



White matter compartment models for *in vivo* diffusion MRI at 300 mT/m



Uran Ferizi^{a,b,c,*}, Torben Schneider^b, Thomas Witzel^d, Lawrence L. Wald^d, Hui Zhang^a,
Claudia A.M. Wheeler-Kingshott^b, Daniel C. Alexander^a

^a Department of Computer Science and Centre for Medical Image Computing, University College London, UK

^b Department of Neuroinflammation, Institute of Neurology, University College London, UK

^c Department of Radiology, New York University School of Medicine, USA

^d A.A. Martinos Center for Biomedical Imaging, Department of Radiology, Harvard Medical School, USA

ARTICLE INFO

Article history:

Received 21 December 2014

Accepted 9 June 2015

Available online 16 June 2015

Keywords:

diffusion MRI

model selection

Connectom

brain microstructure

ABSTRACT

This paper compares a range of compartment models for diffusion MRI data on *in vivo* human acquisitions from a standard 60 mT/m system (Philips 3T *Achieva*) and a unique 300 mT/m system (Siemens *Connectom*). The key aim is to determine whether both systems support broadly the same models or whether the *Connectom* higher gradient system supports significantly more complex models. A single volunteer underwent 8 h of acquisition on each system to provide uniquely wide and dense sampling of the available space of pulsed-gradient spin-echo (PGSE) measurements. We select a set of promising models from the wide set of possible three-compartment models for *in vivo* white matter (WM) that previous work and preliminary experiments suggest as strong candidates, but extend them to fit for compartmental T_2 and diffusivity. We focus on the corpus callosum where the WM fibre architecture is simplest and compare their ability to explain the measured data, using Akaike's information criterion (AIC), and to predict unseen data, using cross-validation. We also compare the stability of parameter estimates in the presence of i) noise, using bootstrapping, and ii) spatial variation, using visual assessment and comparison with anatomical knowledge. Broadly similar models emerge from the AIC and cross-validation experiments in both data sets. Specifically, a three-compartment model consisting of either a Bingham distribution of sticks or a Cylinder for the intracellular compartment, an anisotropic diffusion tensor (DT) model for the extracellular compartment, as well as an isotropic CSF compartment, performs consistently well. However, various other models also perform well and no single model emerges as clear winner. The WM data (with virtually no CSF contamination) do not support compartmental T_2 but partially support compartmental diffusivity. Evaluation of parameter stability favours simpler models than those identified by AIC or cross-validation. They suggest that the level of complexity in models underpinning currently popular microstructure imaging techniques such as NODDI, CHARMED, or ActiveAx, where the number of free parameters is about 4 or 5 rather than 10 or 11, may reflect the level of complexity achievable for a useful technique on current systems, although the 300 mT/m data may support more complex models.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

Magnetic resonance (MR) microstructure imaging uses mathematical models to relate MR signals in each image voxel to microscopic tissue features, and thus estimate and map those features over an image volume. In diffusion MRI, the standard DT model has two key limitations: first it is too simple to explain the data over a wide range of b-values and orientations; and second, it lacks specificity to particular tissue features. A variety of alternative biophysical diffusion MRI models have emerged over the last decade to address these limitations. These models underpin the emerging generation of microstructure imaging techniques

that are now starting to replace DT-imaging in a range of biological and clinical studies into tissue microstructure variation, as in e.g. [Winston et al. \(2014\)](#) and [Kunz et al. \(2014\)](#).

[Stanisz et al. \(1997\)](#) pioneered the multi-compartment representation of separate diffusive processes in nervous tissue. The model has three geometric compartments: ellipsoids for restricted intra-axonal water, anisotropically hindered extracellular water (with diffusivity and relaxation constants different to the intracellular space) and isotropically restricted glial cell water, with exchange between all compartments. The Ball–Stick model ([Behrens et al., 2003](#)) is the simplest two-compartment model that can capture anisotropy. The two compartments exhibit restricted axonal diffusion (through a Stick) and isotropic extra-axonal diffusion (through a Ball). The Composite Hindered and Restricted Model of Diffusion (CHARMED) by [Assaf et al. \(2004\)](#) replaces Behrens' intracellular Stick with impermeable cylindrical fibres, and replaces the

* Corresponding author.

E-mail address: uran.ferizi@med.nyu.edu (U. Ferizi).

extracellular Ball with a full DT. Initially the model used a fixed-diameter distribution to estimate fibre orientation and volume fraction. Later work on AxCaliber (Assaf et al., 2008) additionally estimates the parameters of a gamma distribution of axon diameters (GDR-cylinders). ActiveAx (Alexander et al., 2010; Dyrby et al., 2013) combines elements of Stanisz's model and AxCaliber to obtain the Minimal Model of White Matter Diffusion (MMWMD), which is designed to be the simplest model that adequately fits diffusion MRI data from coherent WM while providing estimates of axon density and diameter. The model has a single axon diameter and cylindrically-symmetric extracellular diffusion, which makes it parsimonious enough to obtain orientationally invariant parameter estimates.

The simple models above do not account for fibre direction inhomogeneity which is abundant in the brain even at a sub-voxel level (Zhang et al., 2011b, 2012; Jeurissen et al., 2013). A wide family of multiple fibre reconstruction algorithms (Seunarine and Alexander, 2009; Tournier et al., 2011) aim to recover multiple fibre orientations or the fibre orientation distribution, but these are not directly relevant here as they do not separate different tissue compartments. However, various compartment models incorporate the idea. Hosey et al. (2005) and Behrens et al. (2007) extend the Ball–Stick model to multiple Sticks, capturing multiple fibre populations with distinct orientation. Similarly, Assaf and Basser (2005) extend CHARMED to multiple intracellular compartments. The DIAMOND model (Scherrer et al., 2013) also employs discrete fibre populations and the number of fibres within each voxel is determined via a model selection framework based on the generalization error. Zhang et al. (2011b) extend the MMWMD, relaxing the assumption of straight parallel fibres to a Watson distribution of orientations. This led to the simpler NODDI model (Zhang et al., 2012), which further assumes zero axon diameter (i.e. Watson distributed Sticks rather than Cylinders). Later work (Tariq et al., 2014) incorporates dispersion anisotropy by replacing the Watson distribution with a Bingham distribution. Concurrently with NODDI, Sotiropoulos et al. (2012) extended the Ball–Stick model to Ball–Rackets also using a Bingham distribution of Sticks, and Bingham mixtures, as in fact proposed earlier in Kaden et al. (2007). Jeurissen et al. (2013) combine the constrained spherical deconvolution model of the fibre orientation distribution from Tournier et al. (2007) with additional grey matter and CSF compartments.

The wide-ranging set of available models relating the diffusion signal to microstructural tissue features raises the question about which are most appropriate in different situations. A series of recent studies have aimed to identify the combination of compartments that best explain the diffusion MRI signal from WM over the accessible range of the measurement space. Panagiotaki et al. (2012) provide a taxonomy of simple compartment models for non-dispersive fibres, by collecting and enriching earlier models. Using the Bayesian Information Criterion (BIC) they then compare the performance in explaining data from fixed rat nerve tissue acquired with a wide range of b-values and diffusion times over long scan times (65 h) and using a high gradient strength (400 mT/m) animal imaging system. Richardson et al. (2013) repeated the study using non-fixed tissue in 'viable' *in vivo* state showing consistency of model ranking but inconsistency of parameter estimates with fixed tissue. Ferizi et al. (2014) extended the acquisition to multi-shell high angular resolution diffusion imaging (HARDI) on a live human volunteer, using a standard human 3T scanner with a maximum gradient strength of 60 mT/m. Each study concludes that, to capture the broad trends in the signal, at least three compartments are required, characterising: i) cylindrically restricted intracellular diffusion; ii) hindered anisotropic extracellular diffusion; and iii) isotropically free and/or restricted diffusion. Another work by Ferizi et al. (2013) extends the set of models to include fibre dispersion via a Watson distribution of fibres and found that, in one of the most coherent WM structures such as the corpus callosum, the dispersed fibre models described the signal better than models with a single orientation. All these experiments show that even in regions of tissue with supposedly the simplest geometry the models required to explain the data are surprisingly complex.

The recent development of human MR systems with 300 mT/m gradients, in particular the *Connectom* scanner (Setsompop et al., 2013), is a major step towards the long-term translation of microstructure imaging techniques to widespread clinical practice. Dyrby et al. (2013) illustrate the benefits of stronger gradients for mapping axon diameters and other parameters in WM using fixed post-mortem tissue and a small-bore animal imaging system. The first experiments verifying those findings on live human subjects are now beginning to emerge from the *Connectom* scanner (McNab et al., 2013; Duval et al., 2014; Huang et al., 2015).

Here we explore the generalisability of earlier model comparison results, which use standard human scanners with up to 60 mT/m gradients, to the wider measurement space of human data accessible with the *Connectom* scanner. We construct a multi-shell HARDI protocol for the *Connectom* scanner with a wide range of b-values and diffusion times. It is similar in spirit and size to that acquired in Ferizi et al. (2014) but exploits the wider measurement space afforded by the 300 mT/m gradients. We concentrate on a subset of models used in previous work (Panagiotaki et al., 2012; Ferizi et al., 2013, 2014) that perform consistently well and compare model rankings obtained from the *Connectom* data sets and data from Ferizi et al. (2014) using the AIC, bootstrapping and cross-validation. We introduce a few methodological enhancements over previous works (Panagiotaki et al., 2012; Ferizi et al., 2013, 2014). First, we fit the models voxel-wise over more homogeneous regions of the genu, midbody and splenium, rather than fit to data averaged across the whole of the corpus callosum. Second, we account explicitly for the variable TE among measurements, by incorporating a compartmental T_2 decay into the modelling. Third, we also study models with compartmentally different diffusivity. We further look at the sensitivity of parameter estimates to noise by looking across bootstrap data sets and to spatial variation by mapping parameters over the corpus callosum, which provide a complementary evaluation of the models.

Methods

This section describes the data acquisition and the identification of regions of interest within the corpus callosum in which we run the model comparison. It then defines the set of models to compare, which enhance earlier models by including both T_2 and compartmental diffusivity effects, and reduce the full set to simplify the comparison. Finally, it describes the methods we use for parameter estimation and model comparison.

Data acquisition

We describe below the acquisition protocols, which are also summarised in Table 1. Ethical approval and written consent was obtained prior to scanning.

The Achieva protocol

This protocol is as described in Ferizi et al. (2014). A PGSE sequence was used on a clinical 3T Philips *Achieva* system, with cardiac gating and TR = 4 s. Every one of the 32 HARDI shells has a unique set of 45 directions with three interwoven $b = 0$ acquisitions. Each shell has a combination of $\Delta = \{30, 50, 70, 90\}$ ms, $\delta = \{6, 10, 15, 22\}$ ms, and $|G| = \{55, 60\}$ mT/m, giving 1,536 measurements and achieving a b-value range of 218 to 10,308 s/mm². TE varied from 51 to 127 ms. The SENSE factor was 1.10.

To cover this protocol, one healthy volunteer is scanned in two separate non-stop sessions, each lasting about 4 h. The nine slices, centred on the mid-sagittal plane, are 4 mm-thick with 2 mm × 2 mm in-plane resolution. The SNR of $b = 0$ images varies from about 30 at TE = 51 ms to 5 for TE = 127 ms. The images are then co-registered using *Flirt* (Jenkinson et al., 2002).

Table 1
The two scanning protocols used, each acquired in ~8 h over two non-stop sessions. The *Achieva* protocol, on the top, contains 32 HARDI shells with 45 unique gradient directions. The *Connectom* protocol, on the bottom, has 48 shells, each with 45 unique gradient directions ('blip-up–blip-down').

<i>Achieva protocol</i>													
	Δ (ms)	Nr	TE (ms)	$ \mathbf{G} $ (mT/m)	b (s/mm ²)		Δ (ms)	Nr	TE (ms)	$ \mathbf{G} $ (mT/m)	b (s/mm ²)		
$\delta = 6$ ms	30	1	51	55	218	$\delta = 15$ ms	30	17	60	55	1,218		
	30	2	51	60	260		30	18	60	60	1,449		
	50	3	71	55	374		50	19	80	55	2,192		
	50	4	71	60	445		50	20	80	60	2,608		
	70	5	91	55	530		70	21	100	55	3,166		
	70	6	91	60	631		70	22	100	60	3,768		
	90	7	111	55	686		90	23	120	55	4,140		
	90	8	111	60	816		90	24	120	60	4,927		
	$\delta = 10$ ms	30	9	55	55		557	$\delta = 22$ ms	30	25	71	55	2,375
		30	10	55	60		687		30	26	71	60	2,868
50		11	75	55	1,010	50	27		87	55	4,470		
50		12	75	60	1,202	50	28		87	60	5,320		
70		13	95	55	1,443	70	29		107	55	6,566		
70		14	95	60	1,718	70	30		107	60	7,814		
90		15	115	55	1,876	90	31		127	55	8,661		
90		16	115	60	2,233	90	32		127	60	10,308		
<i>Connectom protocol</i>													
		Δ (ms)	Nr	TE (ms)	$ \mathbf{G} $ (mT/m)	b (s/mm ²)			Δ (ms)	Nr	TE (ms)	$ \mathbf{G} $ (mT/m)	b (s/mm ²)
$\delta = 3$ ms	22	1	49	61	50	$\delta = 8$ ms	22	25	58	58	300		
	22	2	49	86	100		22	26	58	95	800		
	22	3	49	192	500		22	27	58	190	3,200		
	22	4	49	285	1,100		22	28	58	275	6,700		
	40	5	67	63	100		40	29	72	59	600		
	40	6	67	100	250		40	30	72	100	1,700		
	40	7	67	200	1,000		40	31	72	200	6,850		
	40	8	67	289	2,100		40	32	72	292	14,550		
	60	9	87	63	150		60	33	92	34	300		
	60	10	87	103	400		60	34	92	100	2,650		
	60	11	87	199	1,500		60	35	92	200	10,500		
	60	12	87	290	3,200		60	36	92	292	22,350		
	80	13	107	63	200		80	37	112	61	1,300		
	80	14	107	99	500		80	38	112	100	3,550		
	80	15	107	201	2,050		80	39	112	200	14,150		
	80	16	107	291	4,300		80	40	112	292	30,200		
	100	17	127	63	250		100	41	132	60	1,600		
	100	18	127	101	650		100	42	132	100	4,450		
	100	19	127	200	2,550		100	43	132	200	17,850		
	100	20	127	291	5,400		100	44	132	292	38,050		
	120	21	147	63	300		120	45	152	60	1,950		
	120	22	147	99	750		120	46	152	100	5,350		
	120	13	147	199	3,050		120	47	152	200	21,500		
	120	24	147	291	6,500		120	48	152	292	45,900		

The *Connectom* protocol

This acquisition used the MGH/Harvard-UCLA *Magnetom Skyra Connectom* (Siemens Healthcare) scanner, which has a novel AS302 gradient system with a custom-built 64-channel coil, capable of $|\mathbf{G}| = 300$ mT/m and a slew rate of 200 T/m/s. A PGSE (Tanner and Stejskal, 1968) sequence was used, with GRAPPA parallel imaging, an acceleration factor of 2, cardiac gating and TR = 1 s. The protocol contains 48 HARDI shells, each with 90 directions (45 unique pairs of opposite directions) and 10 interwoven $b = 0$ acquisitions. Each shell has a unique combination of $\Delta = \{22, 40, 60, 80, 100, 120\}$ ms, $\delta = \{3, 8\}$ ms, and $|\mathbf{G}_{max}| = \{60, 100, 200, 300\}$ mT/m (as can be seen in Table 1, most of the actual $|\mathbf{G}|$ values used in the scanning are very close to $|\mathbf{G}_{max}|$ so, for simplicity, we will make no distinction between the two). The b -value therefore ranges from 50 to 45,900 s/mm². Each shell uses the minimum TE possible for a given combination of δ and Δ ; TE thus ranges from 49 to 152 ms. (We note that the difference between the echo spacing, and thus echo train length, for the *Connectom* scanner and conventional scanners, such as *Achieva*, at the 2 mm resolution used, is relatively minor; for our protocol, they were 0.53 ms and 0.72 ms respectively. In addition, the 34.2 mT/m gradient for the EPI

readout in this *Connectom* acquisition was not particularly high, compared with *Achieva's* 35 mT/m, and so any potential Maxwell terms would be negligible.)

For this protocol, the same healthy control as in the *Achieva* acquisition is scanned over two 4 h non-stop sessions. The imaged volume comprises twenty 4 mm-thick whole-brain sagittal slices covering the corpus callosum left-right. The image size is 110 × 110 and the in-plane resolution is 2 mm × 2 mm. The SNR of $b = 0$ images is about 35 at TE = 49 ms and 6 at TE = 152 ms. All *Connectom* images are corrected for eddy current distortions and co-registered using FSL *Flirt/Fnirt* (Jenkinson et al., 2002) and the *eddy* module (www.fmrib.ox.ac.uk/fsl/eddy).

Data sets

Three neighbourhoods of four voxels were selected from the middle of the genu, midbody and splenium in the mid-sagittal corpus callosum. The selection was based on fitting Ball–Stick–CSF model to the entire corpus callosum region, and then selecting from each region four voxels with the highest Stick volume fraction and Stick orientation most closely aligned with the direction perpendicular to the mid-sagittal

slice. We also checked visually that the selected voxels were well-embedded in WM and away from partial volume areas.

We considered various subsets of measurements from these regions of interest to construct different data sets for model fitting and comparison. The following four representative regions of interest were defined on the *Achieva* acquisition:

- *ACH-genu* has four voxels in the centre of the mid-sagittal genu volume;
- *ACH-midbody* similarly contains four mid-sagittal midbody voxels;
- *ACH-splenium* has four mid-sagittal splenium voxels; and
- *ACH-csf* has four mid-sagittal ventricular CSF voxels (used to estimate its T_2).

Four similar regions from the *Connectom* acquisition were defined:

- *CON-genu* has four mid-sagittal genu voxels;
- *CON-midbody* has four mid-sagittal midbody voxels;
- *CON-splenium* has four mid-sagittal splenium voxels; and
- *CON-csf* has four mid-sagittal ventricular CSF voxels.

We also consider several subsets of the *CON-genu* signal to explore the influence of measurements with different gradient strengths:

- *CON-genu-G60* retains only shells with $|G| \approx 60$ mT/m;
- *CON-genu-G100* retains only shells with $|G| \approx 100$ mT/m;
- *CON-genu-G200* retains only shells with $|G| \approx 200$ mT/m; and
- *CON-genu-G300* retains only shells with $|G| \approx 300$ mT/m.

Compartment models

We considered only three-compartment models following earlier findings in Panagiotaki et al. (2012) and Ferizi et al. (2013, 2014). The compartments are shown in Fig. 1, and a more detailed description is given in the Appendix A. The full set of candidate models for each compartment leads to a very large set of three-compartment models. We focus here on a small subset that this earlier work suggests are the strongest candidates. In particular, following Alexander et al. (2010) and Zhang et al. (2012), we consider only isotropic free diffusion for the third compartment, modelling CSF contribution, since we consider only *in vivo*

data. For the intracellular compartments, we consider only Cylinder, Bingham-sticks, and one/two Sticks (we regard the Stick–Stick combination as a single intracellular compartment). Preliminary work suggests possible benefits of Bingham over Watson models of dispersion. However, those experiments reveal that models which contain a more complex compartment than a single-radius Cylinder, such as GDR-cylinders (Assaf et al., 2008; Panagiotaki et al., 2012), dispersed Cylinders (Zhang et al., 2011b) or multiple distinct Cylinders (Zhang et al., 2011a), show only a marginal fitting advantage, and that they also take a very long time to fit and are quite unstable. So we decided to not include these complex models, to keep the set small and the results more comprehensible. Similarly, we consider only Ball and full Tensor for the extracellular compartment, excluding the Zeppelin models to keep the set compact.

Much previous work (Alexander, 2008; Alexander et al., 2010; Zhang et al., 2011b, 2012; Panagiotaki et al., 2012; Ferizi et al., 2013, 2014) assumes equal diffusivity of the material in the intracellular and extracellular compartments. For each of our compartment models, we evaluate performance with and without this constraint. To highlight the difference in later results, the models with separate intra/extracellular diffusivities have a “-diff” suffix.

In the model comparison we add a combination of two cylindrically symmetric tensors, Zeppelins, with unrelated orientations. This type of model, often referred to as a mixture of Gaussians, has been used to recover intra-voxel populations of fibres (Tuch et al., 2002; Parker and Alexander, 2003). We also further consider two established models from the earlier literature, NODDI and MMWMD as implemented similarly in Zhang et al. (2012) and Alexander et al. (2010). Both models use a simple tortuosity approximation, which the other models we consider here do not. They also have fixed diffusivity parameters, here to 1.9 and $1.7 \mu\text{m}^2/\text{s}$ respectively.

T₂ effect

Varying TE means that we need to model T_2 effects. The total signal model is then:

$$S = S_0 \left(f_i \exp\left(-\frac{TE}{T_2^i}\right) S_i + f_e \exp\left(-\frac{TE}{T_2^e}\right) S_e + f_c \exp\left(-\frac{TE}{T_2^c}\right) S_c \right) \quad (1)$$

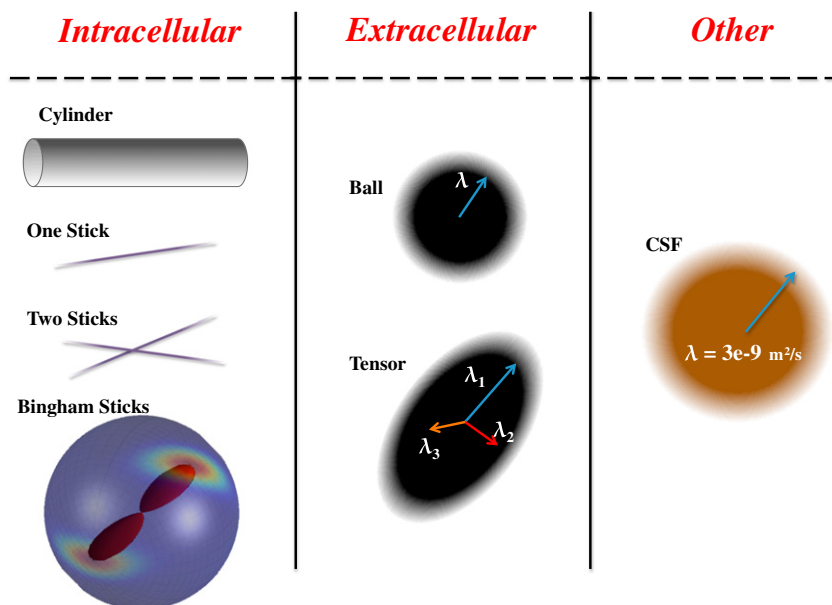


Fig. 1. The elements of each compartment class designed to capture diffusion through a particular tissue medium: intracellular, extracellular, and the CSF. A model consists of a combination of three compartments, one from each class. More details can be found in sub-Section 2.3 and in the Appendix A.

where f_i , f_e and f_c are the weights of the normalised intracellular, extracellular, and CSF compartment signals S_{intra} , S_{extra} and S_c , respectively; the values of compartmental T_2 decay rates are indexed similarly; S_0 is the proton density signal (which is TE-independent, and obtained from fitting to the $b = 0$ signal; see next section).

Model fitting

We use a two-stage model fitting procedure. The first step estimates the T_2 of tissue separately in each voxel by fitting to the (TE-dependent) $b = 0$ signal a bi-exponential model in which one component is from tissue and the other from CSF. A preliminary analysis of voxels fully inside WM regions shows no significant departure from a mono-exponential decay, so we assume equal T_2 within the intracellular and extracellular compartments. When fitting the bi-exponential model, the value of T_2 in CSF is fixed to 1,000 ms. We fix this *a priori* as a more precise value of CSF is unlikely to be estimated with the range of TEs we use; this is also what we approximately estimate from an ROI in the ventricles. Thus, for each voxel we estimate the volume fraction of CSF, the S_0 and the T_2 of the tissue. These three estimates are then fixed for all the subsequent model fits, which enable a more direct comparison of the classes of WM compartments on their ability to fit the diffusion-weighted signal.

In the second step, we fit each model using the Levenberg–Marquardt algorithm with an offset-Gaussian noise model, as in Panagiotaki et al. (2012) and Ferizi et al. (2014). The offset-Gaussian objective function is an approximation to a non-central chi noise model (Gudbjartsson and Patz, 1995), and accounts for the bias in low signal measurements introduced by the noise floor. It minimises the sum-of-squared-errors:

$$SSE = \sum_{i=1}^N \frac{(\tilde{S}_i - \sqrt{S_i^2 + \sigma^2})^2}{\sigma^2} \quad (2)$$

where N is the number of measurements, \tilde{S}_i is the i -th measured signal, S_i its prediction from the model, and σ is the noise standard deviation. To find σ for the *Connectom* data we used the background signal, as well as ROI signal in $b = 0$ images. For *Achieva* data we only used the ROI signal, as these images are acquired with a limited FOV. In both cases the σ estimates matched reasonably well with the observed signal noise floor in the plots, i.e. the residual signal along the fibres at the highest b -value.

Each fitting was repeated 500 times, each time randomly perturbing from a central starting point with equal volume fractions, axial WM diffusivity $2 \mu\text{m}^2/\text{s}$, radial diffusivity $1 \mu\text{m}^2/\text{s}$, and cylinder radius $2 \mu\text{m}$. The CSF diffusivity was fixed throughout to $3 \mu\text{m}^2/\text{s}$.

Model ranking

We compare the models in two ways. First, we use the Akaike information criterion (Akaike, 1974)

$$AIC = -2\log(L) + 2K \quad (3)$$

where, given the data, L is the likelihood of the model with K number of free parameters. It is standard practice to drop the constant terms in $\log(L)$ because the absolute value of AIC is not informative, and that the AIC is not used to compare the performance of models across data sets (Burnham and Anderson, 2002). Further, as the noise level σ is found from the images, thus known *a priori* in the model fit, the first term in the AIC equates to the SSE of Eq.(2). The lower the AIC , the better the model. Compared with other similar studies which use BIC (Panagiotaki et al., 2012; Ferizi et al., 2014), here we switched to the AIC following the recommendation in Burnham and Anderson (2002);

however, for our large data sets the differences between AIC and BIC are minor.

In addition to the AIC , we use cross-validation (Stone, 1974) and bootstrapping (Efron, 1979) as complementary methods for model comparison and to further investigate the stability of the model ranking and parameter estimation. We use leave-one-out cross-validation at the level of HARDI shells, i.e. at each iteration we leave out a complete HARDI shell, fit the model to the remaining data, and use the fitted parameters to estimate all measurements in the missing shell. The final score is the average fitting error over all shells (32 for the *ACH-genu* data and 48 for the *CON-genu* data). When bootstrapping, we sample with replacement in the original data to get the same number of measurements. We bootstrap fifty times for each of the four voxels in the genu, giving a total of 200 data sets to which we fit the models. For our data, that number of bootstraps is sufficient for the variance in the mean parameter estimates over bootstraps to stabilise within 5% of its value.

Each model is fitted separately in each voxel of the regions of interest defined in Section 2.2, and the results reported in Section 3 are, unless otherwise stated, averages over all the voxels in the region.

Results

This section starts with some images of the corpus callosum. Then we show the ranking obtained via various model selection procedures after fitting the models. Finally, we evaluate the parameter estimates and stability obtained for each model.

Connectom data; T_2 fitting

Fig. 2 shows the diffusion-weighted images from the *Connectom* scanner with gradient applied in the direction perpendicular to the fibres of the corpus callosum. Fig. 3 shows the ROI-average diffusion weighted signal for each combination of pulse sequence parameters in the *CON-genu* data set, split into the four gradient strengths used, together containing 4,800 measurements. The measurements provide good coverage of the signal range. Anisotropy is apparent even at the highest diffusion weighting of $45,900 \text{ s/mm}^2$, clearly supporting models that assume restricted diffusion within axons. There were no obvious differences between the 45 pairs of opposite diffusion encoding directions, other than what is expected from measurement noise. (In two shells with $\delta = 3$ and 8 ms , $|G| = 60 \text{ mT/m}$, and $\Delta = 20 \text{ ms}$, four or five pairs showed unexpected differences at around twice the level of noise standard deviation. We ran the model fitting with and without these outliers, and the changes observed in the model parameter estimates were insignificant.)

Fig. 4 shows the distribution of log-signal at $b = 0$ versus echo time TE for WM (genu, midbody and splenium) and CSF. In both WM and CSF the data show no significant departure from the mono-exponential model, so for each scanner we concluded that, at least for these data sets, the intracellular and extracellular T_2 are indistinguishable. Voxel-wise estimated T_2 of *ACH-genu* was $59 \pm 2 \text{ ms}$, and of *CON-genu* was $51 \pm 1 \text{ ms}$. Away from the genu, there was variation across the regions of the corpus callosum. A similar difference can be seen for higher T_2 in the splenium, at $70 \pm 1 \text{ ms}$ and $61 \pm 1 \text{ ms}$ for *ACH-splenium* and *CON-splenium*, respectively; but we found more comparable $63 \pm 3 \text{ ms}$ for *ACH-midbody* and $64 \pm 2 \text{ ms}$ for *CON-midbody*. The T_2 estimate from *ACH-csf* data averaged about $1,500 \pm 400 \text{ ms}$, and $700 \pm 86 \text{ ms}$ from *CON-csf* data, although these estimates are less stable because the slope of the linear fit is very shallow. These estimates are in the order of previous estimates (MacKay et al., 1994; Whittall et al., 1997) of about 71 ms for genu, and over- $1,000 \text{ ms}$ for CSF.

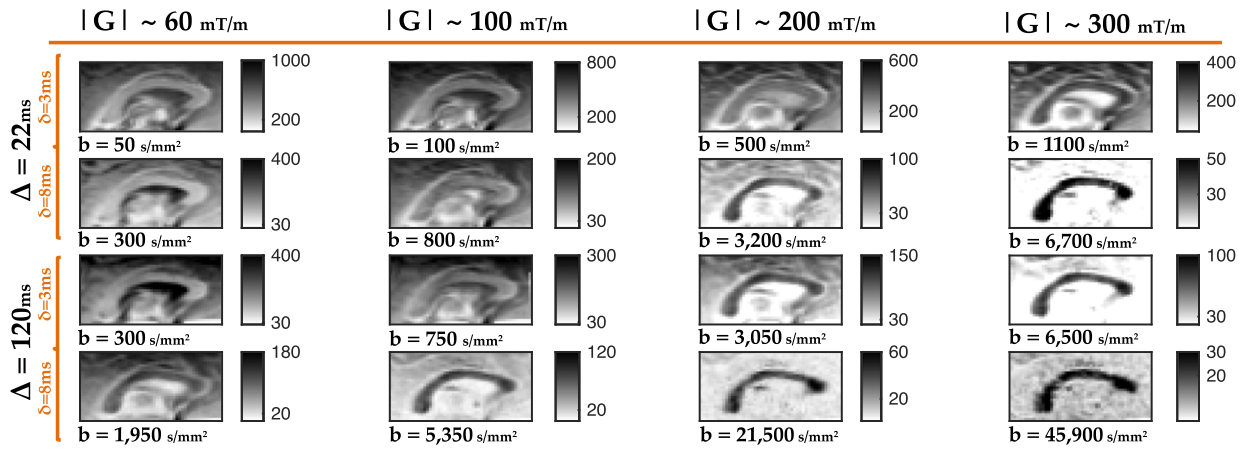


Fig. 2. Images of the corpus callosum showing the signal for gradient direction perpendicular to the fibres, at each $|G|$, for each pulse time δ , but only for the smallest and largest diffusion times Δ . The grey scale is inverted and adjusted in each case so as to give a reasonable contrast between the corpus callosum and the background. Signal still persists even at $b = 45,900 \text{ s/mm}^2$.

Model ranking

Fig. 5 compares AICs of each model for the three regional data sets from both scanners. The scores for each voxel are shown as coloured marks, and the star marker gives the average over the four-voxel region. We can roughly split the models into three groups. The first group of six models, marked with an H on the plots, generally has the lowest AIC; this group includes five models with separate compartmental diffusivities, comprising all four Tensor models as well as the combination of Ball with Bingham, and the remaining Tensor-Bingham model. The models in group marked with L generally have the highest AICs, and includes all combinations of extracellular Ball and intracellular Stick/s. The remaining models form a third group, marked with M on the plot, that have broadly intermediate AIC scores. The division into these groups is present in both genu and splenium data sets. AICs from the midbody region are less consistent, arising largely because of partial volume with CSF. We concentrate

from here onwards on the genu data sets for a more detailed analysis of model performance.

Fig. 6 shows the AICs of the models for the subsets of the CON-genu data sets with different gradient strengths. The plots show clearly how the group structure emerges in the plots as the gradient strength increases; specifically, model combinations of Ball with Cylinder are increasingly favoured, but those with Stick/s are not. Bingham-based models perform best across the four datasets. The results from CON-genu-60 and ACH-genu are largely similar, since both use similar gradient strengths; however, we note that some differences would arise as ACH-genu uses much larger δ , and so has higher b-value measurements than CON-genu-60.

Fig. 7 shows how the ranking of models varies over bootstrap iterations and cross-validation folds. In bootstrapping, at each iteration, we compute the AIC of each model from the seen data; in cross-validation, we compute the SSE of the unseen data. Thus we obtain a model ranking from each iteration of each procedure. The positional

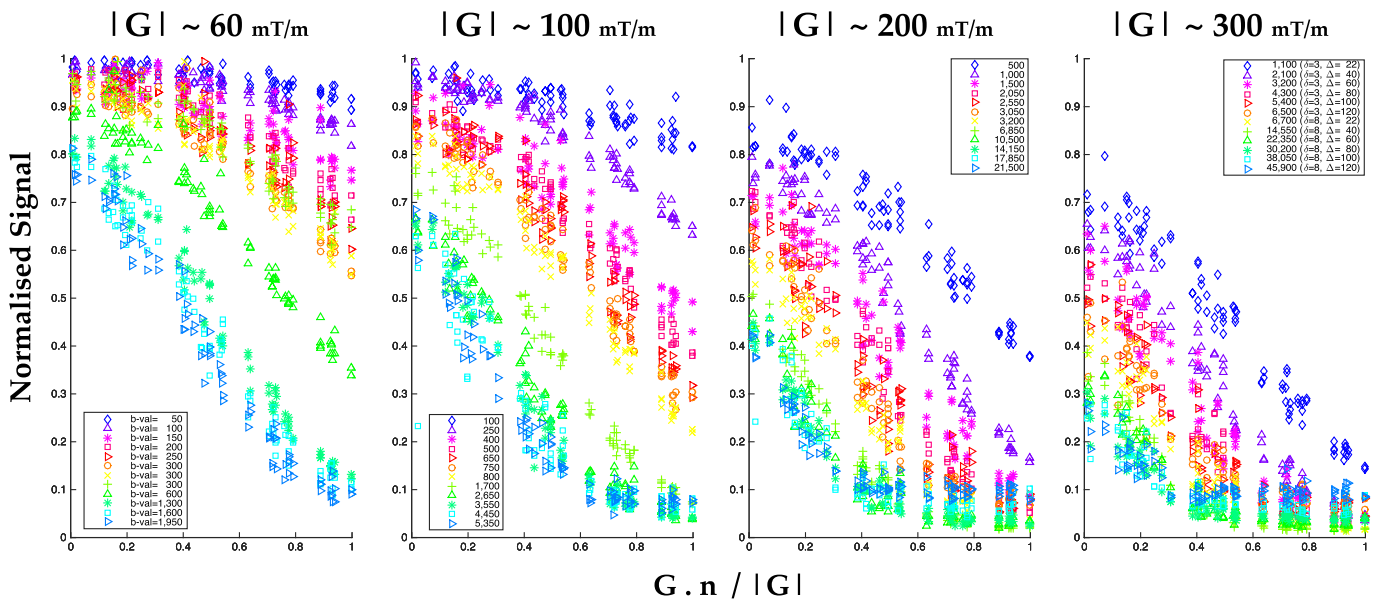


Fig. 3. Diffusion weighted signal in the CON-genu data set averaged over the four voxels in the middle of the genu. For clarity, the signal is split across each gradient strength. The legend gives the b-value in units of s/mm^2 and, in the last plot, also diffusion and pulse times (δ, Δ), which are the same across the four plots, in units of ms. The x-axis gives the cosine of the angle between the applied gradient vector G and the fibre direction \mathbf{n} ; at 0 on the left the gradient is perpendicular to the fibres, but as it approaches 1 on the right the gradient direction becomes parallel with the fibre.

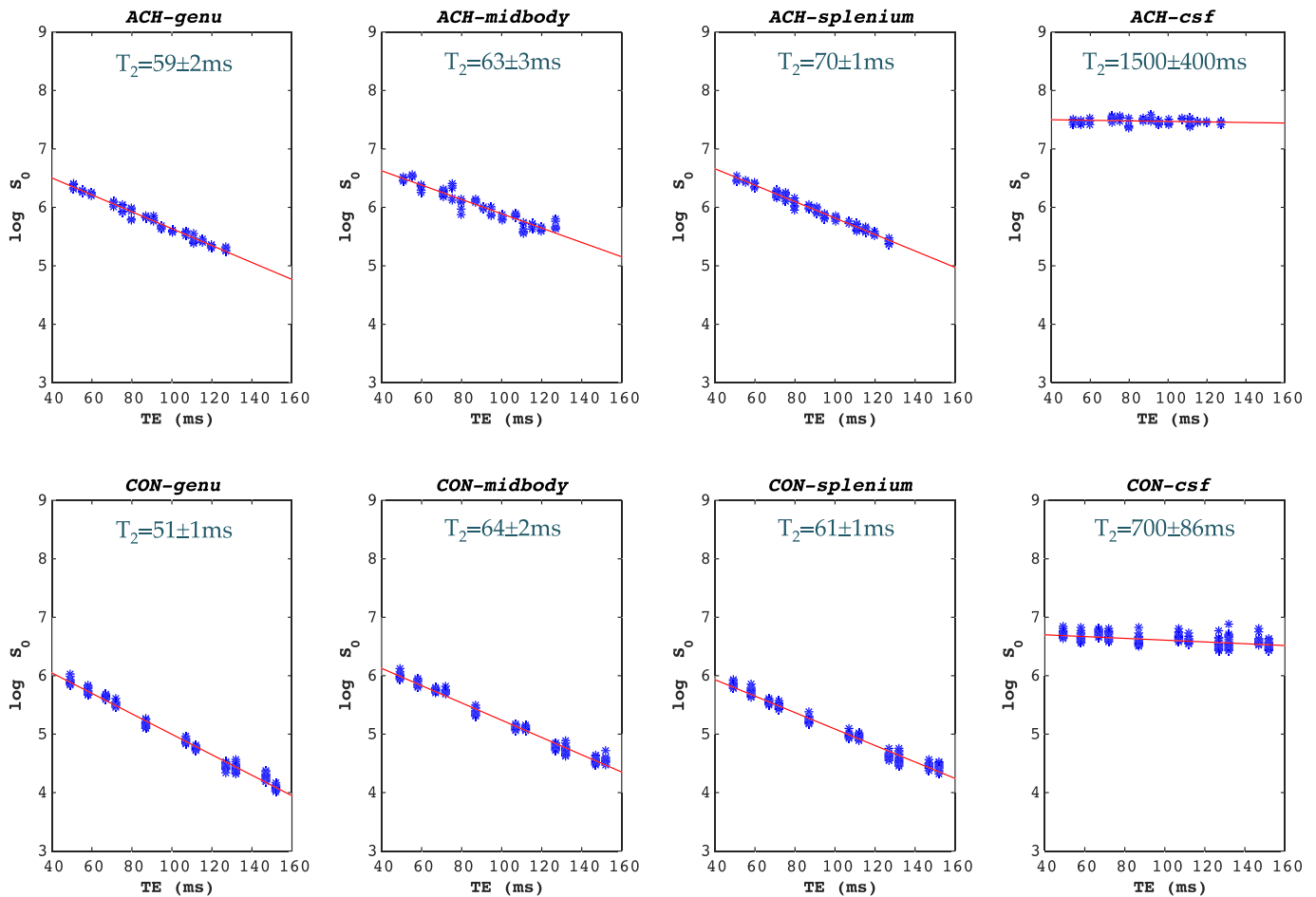


Fig. 4. The blue markers of the plots show the $b = 0$ signals from one-of-four ROI voxels across TE values, as used in the estimation of T_2 ; the T_2 statistics in each subplot are taken across the four ROI voxels. *Achieva* plots are on the top-half, with *Connectom* plots below them; refer to the main text for the names of the data sets. Each of sixteen *Achieva* TEs has six $b = 0$ signals, while the *Connectom* data has forty $b = 0$ signals for each of the twelve TEs. The gradient of the fitted line, shown in red, is an estimate of the negative inverse of the T_2 value (the T_2 values inside each plot are estimates extracted from a bi-exponential fit, however, which also includes a CSF component). The distribution of points in WM did not suggest two separate rates of T_2 decay, therefore we fixed the intra/extracellular T_2 of WM compartments to be the same.

variance diagrams show a histogram for each model of the position in the ranking over all iterations, pooled over the four genu voxels. For example, the top left diagram shows that Tensor–Bingham–CSF–diff and Tensor–Cylinder–CSF–diff consistently give the lowest AIC, whereas the bottom left diagram shows that their ranking by SSE on the unseen shells has lower rank more often.

The bootstrap and cross-validation results broadly reflect the group structure observed from AIC in Fig. 5. The leave-one-shell-out strategy in our cross validation creates much greater variation in the model ranking than the random selection of measurements in the experiment that does not consider the shell structure of the acquisition scheme (results not shown). This shows that the full distribution of b -values is much more influential on the choice of model than the choice of gradient directions.

There is a similar picture of model ranking stability when using *CON-genu* data. The group structure remains in bootstrapping and cross-validation, though there are more variations within the groups.

Fig. 8 illustrates the fit of three models to the raw signal from just one of the voxels from the four-voxel neighbourhoods contributing to *ACH-genu* and *CON-genu*; the results are similar in the other voxels. We pick a representative from each group in the ranking of *ACH-genu* in Fig. 5: Tensor–Bingham–CSF–diff (top performing group, H), Ball–Bingham–CSF (*M* group), and Ball–Stick–CSF (*L* group). The model signal is shown as solid line, whereas the raw data is shown

with markers. We can see that all three models capture the *ACH-genu* data reasonably well. The plots on the right reveal that higher b -value shells of *CON-genu* data enable a greater differentiation between the models. In particular, the three top-right plots reveal that the simpler Ball–Stick–CSF is clearly worse at capturing the 5,400 and 6,500 s/mm^2 shells. This pattern continues with even higher b -value shells in the bottom panel of plots. Beyond the 14,550 s/mm^2 shell we see differences between the two Bingham models: while Ball–Bingham–CSF cannot capture larger b -value shells, the extra complexity of Tensor–Bingham–CSF–diff model allows it to capture higher b -value shells more closely.

Parameter values and stability

Table 2 compares parameter estimates from each model from the voxels in *ACH-genu* (top) and *CON-genu* (bottom). Each estimate is the mean obtained after fitting models separately in each of the four voxels. Most models fitted to *ACH-genu* data give higher intracellular than extracellular volume fraction, which is consistent with what we expect from histological studies of WM tissue (Aboitiz et al., 1992; Highley et al., 1999); however, combinations of Tensor with Cylinder or Stick give lower intracellular volume fraction estimates. The coupled diffusivities are highest in Bingham models, which makes sense because they can explain reduced apparent diffusivity by fibre dispersion. In models

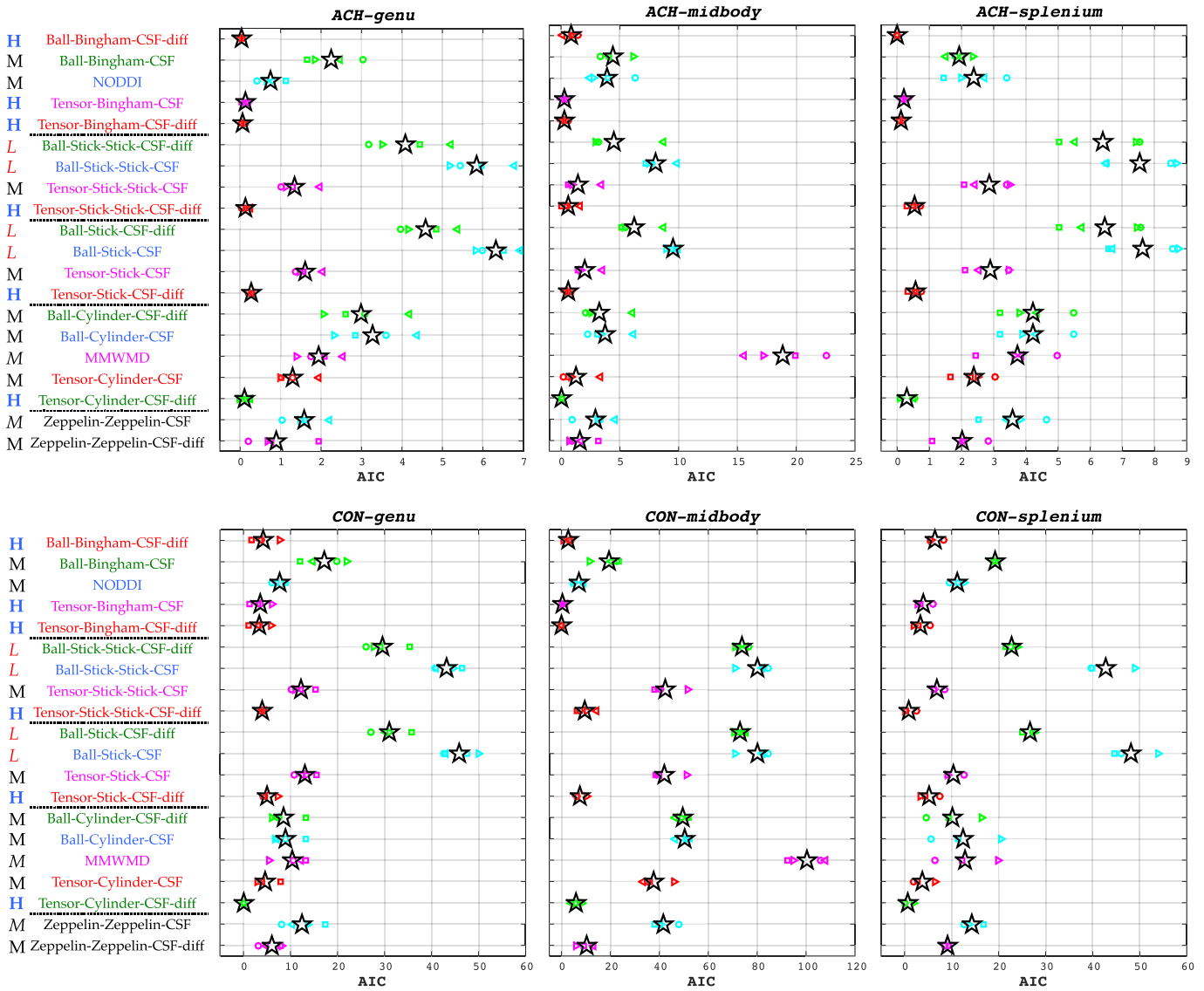


Fig. 5. Model ranking for each data set. Three model clusters are identified. On top of the ranking are six models, five of which have separate compartmental diffusivities (all four Tensor models, and Ball–Bingham), and the remaining Tensor–Bingham model; these are marked with H beside the plots. Combinations of Ball with Stick/s, marked with L, rank last. Middle-ranking models are marked with M. The colours are simply to discriminate rows and the coloured markers, one for each component voxel, correspond to the coloured name to the left. Each voxel has a different marker shape, which is consistent from row to row. The black stars show the mean over the voxels in each row. AIC values are meaningful within each data set, but not across the data sets.

with separate compartmental diffusivities, the intracellular diffusivity is generally higher than the extracellular diffusivity in Ball models, but the opposite applies to Tensor models (except Tensor–Bingham–CSF-diff). The Bingham distribution is more dispersed in Ball than Tensor models, most likely because the model has to compensate for the lack of extracellular anisotropy.

We see similar patterns in volume fraction, dispersion and diffusivity for the models fitted to *CON-genu* data. Again, separating compartmental diffusivities generally makes the volume fractions closer.

The parameter values from simple mixture models are revealing to compare. Specifically, we see in Stick–Stick and Zeppelin–Zeppelin models which provide more angular flexibility that the second tissue compartment, of small volume size in the case of Stick, is almost perpendicular to the first. Rather than capturing intrinsic anatomical features this is a sign that these models fail to capture the signal variation that arises from dispersion, and so end up with unexpected configurations simply to capture that effect, as illustrated in Jbabdi et al. (2012).

Parameter estimates from *ACH-genu* and *CON-genu* are broadly consistent, especially for simpler models. For these simpler models, the radius estimates are generally higher in *ACH-genu*. This suggests that the higher gradient strengths of *CON-genu* help ameliorate the overestimation of axon diameter index, consistent with Dyrby et al. (2013).

Table 3 summarizes results using bootstrap samples to study the variance of the different parameter estimates. Comparison with Table 2 shows that the estimates from the full data set are reasonably similar to the mean across the bootstrap data sets. In addition, simpler models tend to have higher stability in the parameter estimation, i.e. the parameter estimates have lower variance across bootstrap samples.

Fig. 9 shows the spatial variation over the corpus callosum of parameter estimates obtained in four different models. Specifically, we pick two relatively simple models representing established techniques in the literature, MMWMD and NODDI, and two more complex models, Tensor–Cylinder–CSF-diff and Tensor–Bingham–CSF-diff, that have similar combinations of components but consistently outperform the simpler models in our model comparison experiments. The key

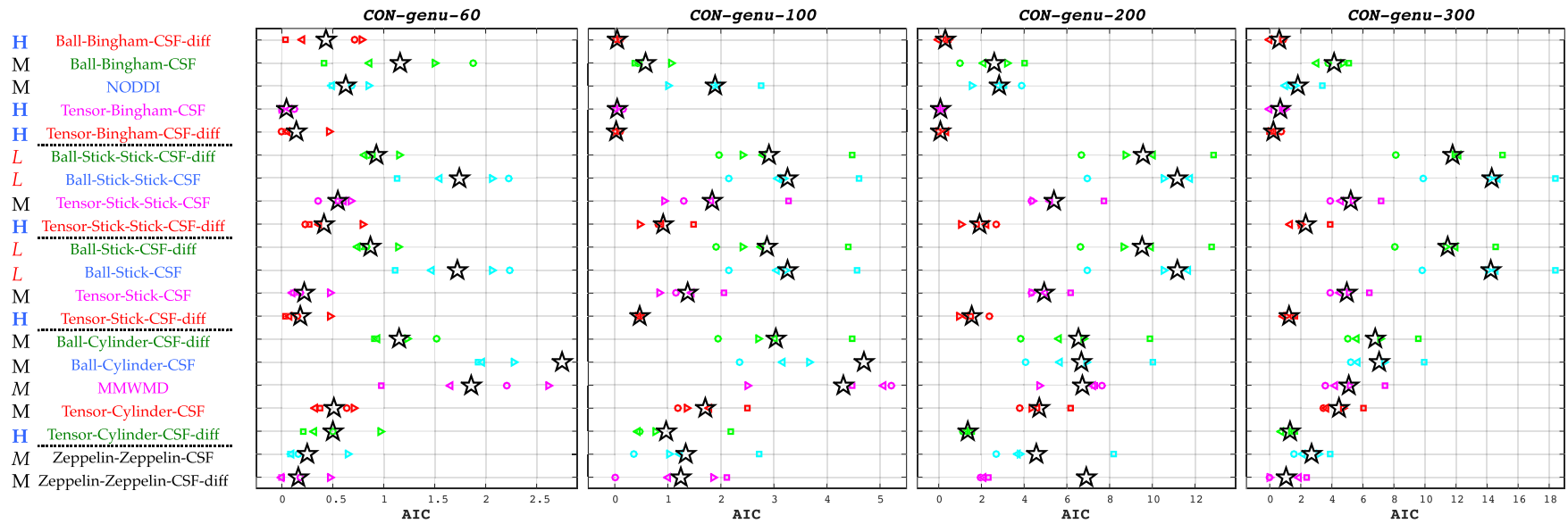


Fig. 6. The stability of model ranking with gradient strength. All four gradient-specific data sets reveal clear differences between model classes. Specifically, Ball with Cylinder models are favoured by higher gradients/b-values, whereas Ball with Stick/s combinations are not. *CON-genu-60* uses similar gradient strengths as *ACH-genu* but it has a lower $b_{max} \sim 2,000 \text{ s/mm}^2$ vs. $b_{max} \sim 10,000 \text{ s/mm}^2$.

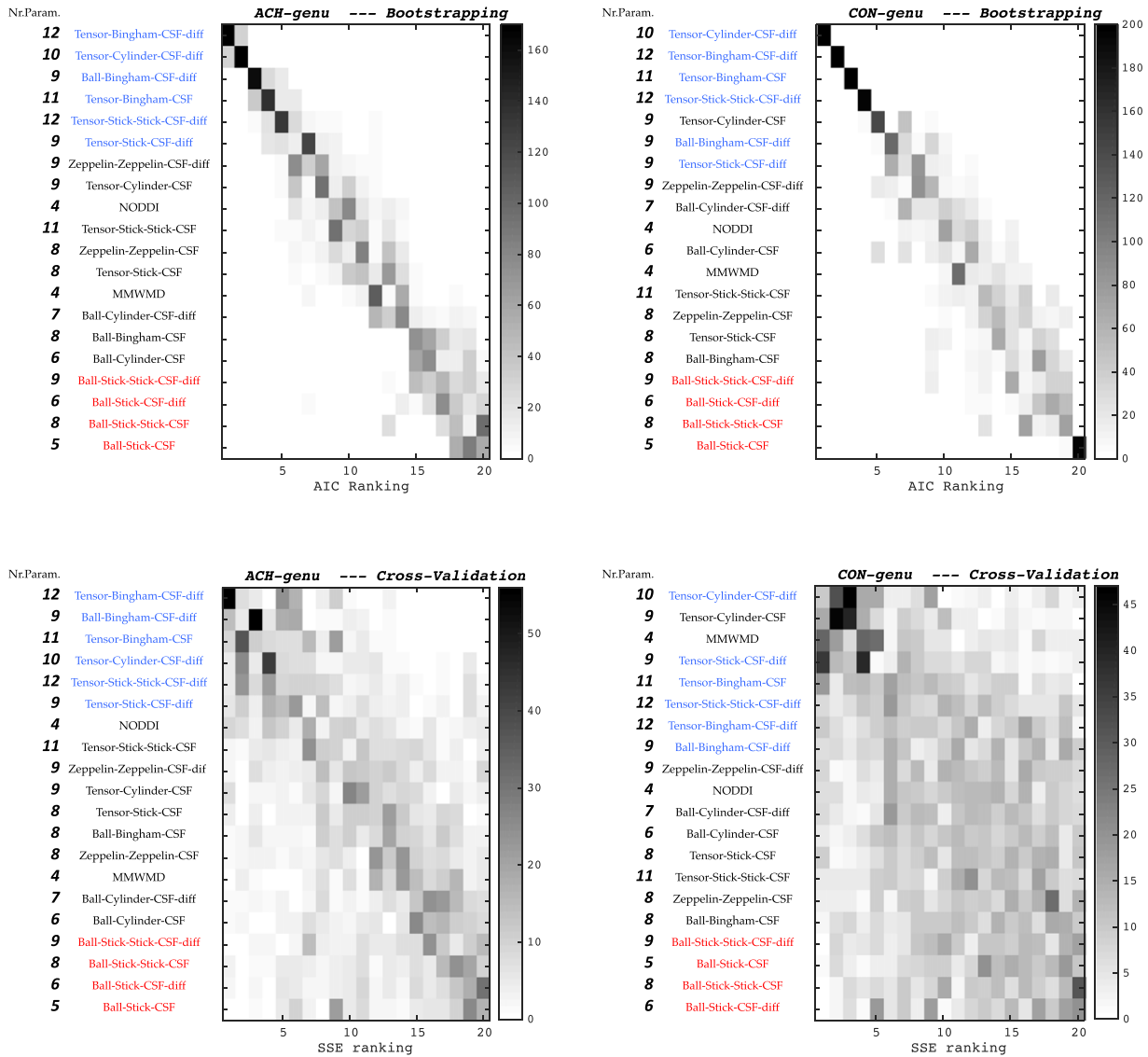


Fig. 7. The AIC of bootstrap data sets (top) and SSE of cross-validation data sets (bottom) from the *ACH-genu* and *CON-genu* (see sub-Section 2.6 for more details). In bootstrapping we construct 200 data sets, fifty for each of the four voxels in the ROI. In cross-validation, each HARDI shell is left aside at a time. Clearly, there is greater uncertainty when predicting unseen shells in the cross-validation, especially for the richer *CON-genu* data set. Generally, in all diagrams the six H-ranked models, shown in blue, are on top, while the red-marked L-models are last. The number of model parameters are shown on the left (excluding three, S_0 , CSF volume, and T_2 of WM).

observation is that the simpler models produce smoother parameter maps that are more consistent with trends in volume fraction that we expect to observe. In particular, the parameter maps from the complex models fitted to the *Achieva* data set are noisy whereas those from the simpler models are much smoother. The *Connectom* data appears to support the complex models better and produces more qualitatively acceptable parameter maps; the variation is still higher than for the simple models but it may reflect genuine anatomical differences. These maps emphasise that fitting data and predicting unseen data better are not the only important criteria in selecting an appropriate model for parameter mapping, as simpler models offer greater stability and so may reflect a more correct sensitivity to pathology or underlying anatomical differences.

Discussion

In this work we sampled a wide measurement space of *in vivo* human data accessible with the *Connectom* scanner and its 300 mT/m gradients with the aim of determining whether similar compartment

models for WM explain these unique data as more accessible data from standard systems with 60 mT/m. Thus, we compare various models using acquisitions that span the range of possible PGSE measurements as widely and densely as possible from both the *Connectom* and a standard *Achieva* scanner. We compare models in several ways: AIC and cross-validation to determine their ability to explain measured data and predict unseen data, parameter variance across bootstrap samples to assess stability in the presence of noise, and spatial variation of parameter estimates to assess stability more qualitatively and to check for recovery of expected trends in the corpus callosum.

First, we observe (through Figs. 2 and 3) clear signal and signal anisotropy in *Connectom* data even with b-values as high as 45,900 s/mm², which shows very strong diffusion restriction. We found (in Fig. 4) that neither data set supports separate T_2 parameters for the intracellular and extracellular compartments, but they do support a separate T_2 for CSF. The estimate of T_2 varies across the corpus callosum, being lower in the genu than in the splenium and the midbody. We see T_2 differences across the two data sets, but the trends are similar: *Achieva* T_2 values are about 10 ms higher in both the genu and splenium regions.

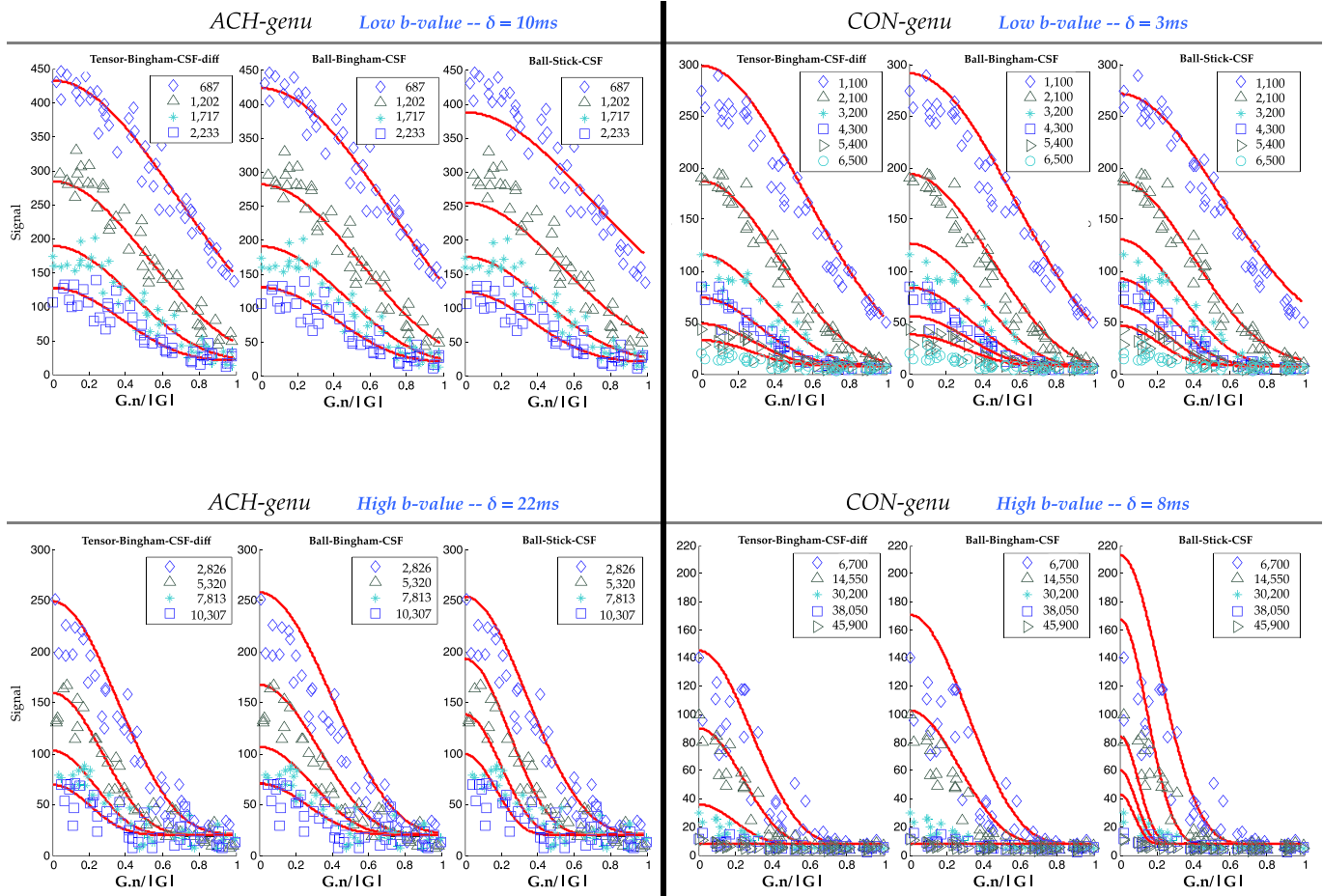


Fig. 8. Plots illustrating the quality of fit for three selected models fitted to *ACH-genu* and *CON-genu*. We select one voxel from the genu ROI, whose raw signal is shown with markers and the model signal shown as solid line. Though the models are fitted to all the data, for clarity, plots show only a few selected high and low b -value shells, of $|G| = 60$ mT/m for *ACH-genu*, and 300 mT/m for *CON-genu*, leaving out $b = 0$ measurements. The shells have the same Δ (30, 50, 70, 90 ms for *ACH-genu*, and 22, 40, 80, 100, 120 ms for *CON-genu*), increasing in b -value from top to bottom. While the Ball–Stick model clearly fits the signal worse, differences between the two Bingham models are more subtle. However, the Tensor model has extra flexibility to capture signals across the full range, e.g. matching the higher b -value shells in *CON-genu* better than the Ball–Bingham model.

Though T_2 values are in the order of previous estimates, slight variations can be expected compared with other experiments specifically designed to measure tissue T_2 . The fixed 1,000 ms T_2 of CSF in both datasets may slightly affect WM T_2 . There are differences in the EPI-train and bandwidth in the two acquisitions that may contribute to the differences in observed T_2 , as may the T_2^* effects. T_1 relaxation can potentially affect the T_2 estimate, but the linear fit in Fig. 4 suggests that this effect is negligible (which is to be expected from a relatively high white matter $T_1 \sim 800$ ms, and a higher density of shells at lower TEs). However, we believe the dominant effect is partial volume contamination with CSF, particularly in the midbody, which differ in the two acquisitions.

We also found (from Figs. 5 and 7) that a broadly consistent group of three-compartment models emerge as best at explaining both the *Achieva* and *Connectom* data from the genu. Specifically, Tensor–Bingham or Tensor–Cylinder combinations perform strongly, using both *AIC* and cross-validation, with Tensor–Stick–Stick and Ball–Bingham also ranking consistently highly. The high performing models give a qualitatively reasonable fit across the range of b -values with no obvious sign of overfitting (Fig. 8). Moreover, we see (in Fig. 6) that higher gradients data favour more the complex fibre configurations when compared with the relatively simple Stick combinations.

The highest performing models, in particular Tensor–Bingham, benefit only slightly from having separate intracellular and extracellular diffusivity parameters. The separation appears more beneficial for models with Stick and Cylinder intracellular compartment, which

may suggest that the performance enhancement simply comes from the additional parameter enabling these models to capture signal variation arising from orientation dispersion. Both midbody data sets have more partial volume artefacts, and provide a less clear comparison, although the results broadly support those from the genu and splenium.

The analysis of parameter stability (Tables 2 and 3, and Fig. 9) favours simpler models. First, the extra parameter introduced by separating intracellular and extracellular diffusivities makes the resulting volume fraction parameter estimates less in line with what we expect from histology, especially in Tensor models. Furthermore, separating the diffusivity parameters often increases the variance of diffusivity parameter estimates over bootstrap samples. Parameter maps over the corpus callosum appear noisier for the complex models emerging from the *AIC* and cross-validation rankings than simpler models underpinning standard techniques, such as NODDI and ActiveAx/MMWMD, although the maps from more complex models appear less noisy in the 300 mT/m *Connectom* data than the 60 mT/m *Achieva* data.

Our results suggest two broad conclusions about the choice of models for microstructure imaging. First, the choice of models to underpin microstructure imaging applications goes beyond the ability of the models to explain and predict data, as assessed by *AIC* and cross-validation. While these quantities are informative to evaluate, and fitting and predicting data are clearly desirable qualities, a useful microstructure imaging technique must also provide realistic models, with

Table 2

Parameter estimates from model fitting to the full data set of *ACH-genu* (top) and *CON-genu*. The CSF volume fraction and T_2 are the same across all models. κ_1 gives the degree of fibre dispersion and κ_2 the extent of fanning. Angle₁ refers to the angle, in degrees, between the first fibre compartment and the direction perpendicular to the mid-sagittal slice; similarly, Angle₂ corresponds to the second fibre compartment (the 2nd Stick or Zeppelin). In each data set, the SSE scores are projected onto a 0 (best fit) to 100 (worst fit) scale.

		Cylinder/Bingham/Stick						Tensor/ Ball							
Models		SSE (0=best)	Vol Fraction	Diffusivity ($\mu\text{m}^2/\text{s}$)	κ_1	κ_2	Radius (μm)	Angle ₁ (°)	Angle ₂ (°)	Vol Fraction	Axial Diff. ($\mu\text{m}^2/\text{s}$)	Radial Diff. ($\mu\text{m}^2/\text{s}$)	Radial Diff. ($\mu\text{m}^2/\text{s}$)		
H	Ball-Bingham-CSF-diff	1.1	0.74	2.66	14	8		0.9		0.24	0.71			ACH-genu	CSF Vol.Fract. = 0.02
M	Ball-Bingham-CSF	36.3	0.84	2.21	10	5		0.9		0.14	2.21				
M	NODDI	13.9	0.63	1.90	11	0		0.9		0.35	1.90	0.67			
H	Tensor-Bingham-CSF	2.0	0.63	1.96	22	16		0.9		0.35	1.96	0.66	0.57		
H	Tensor-Bingham-CSF-diff	0.0	0.71	2.62	15	9		0.9		0.27	0.93	0.67	0.58		
L	Ball-Stick-Stick-CSF-diff	63.7	0.58	2.01				0.9	83.5	0.38	0.90				
L	Ball-Stick-Stick-CSF	93.3	0.62	1.53				0.9	81.8	0.33	1.53				
M	Tensor-Stick-Stick-CSF	21.9	0.45	1.68				0.9	82.4	0.51	1.68	0.64	0.46		
H	Tensor-Stick-Stick-CSF-diff	2.7	0.39	1.00				0.9	83.9	0.57	2.58	0.56	0.36		
L	Ball-Stick-CSF-diff	72.7	0.59	2.00				0.9		0.39	0.90				
L	Ball-Stick-CSF	100.0	0.63	1.50				0.9		0.34	1.50				
M	Tensor-Stick-CSF	25.2	0.46	1.66				0.9		0.52	1.66	0.64	0.44		
H	Tensor-Stick-CSF-diff	3.8	0.40	0.98				0.9		0.58	2.59	0.57	0.38		
M	Ball-Cylinder-CSF-diff	48.0	0.78	1.94			6.2	0.9		0.20	0.61				
M	Ball-Cylinder-CSF	53.1	0.84	1.61			6.1	0.9		0.14	1.61				
M	MMWMD	32.6	0.63	1.70			5.0	0.9		0.35	1.70	0.60			
M	Tensor-Cylinder-CSF	20.7	0.44	1.66			6.9	0.9		0.54	1.66	0.40	0.27		
H	Tensor-Cylinder-CSF-diff	1.9	0.47	1.04			3.5	0.9		0.51	2.75	0.57	0.35		
M	Zeppelin-Zeppelin-CSF	25.5	0.56	1.74				0.4	2.0	0.42	1.74	0.83			
M	Zeppelin-Zeppelin-CSF-diff	12.0	0.52	1.96				1.0	89.6	0.46	1.99	0.33			
WM $T_2 = 59$ ms															
H	Ball-Bingham-CSF-diff	18.2	0.77	2.35	12	5		0.1		0.22	0.96			CON-genu	CSF Vol.Fract. = 0.01
M	Ball-Bingham-CSF	40.8	0.84	2.16	12	7		0.1		0.15	2.16				
M	NODDI	21.9	0.64	1.90	13	0		0.1		0.35	1.90	0.67			
H	Tensor-Bingham-CSF	13.2	0.64	1.95	20	12		0.1		0.35	1.95	0.69	0.56		
H	Tensor-Bingham-CSF-diff	12.4	0.57	1.59	23	13		0.1		0.42	2.46	0.63	0.47		
L	Ball-Stick-Stick-CSF-diff	64.9	0.60	2.10				0.1	26.5	0.38	0.85				
L	Ball-Stick-Stick-CSF	93.2	0.64	1.54				0.1	79.0	0.34	1.54				
M	Tensor-Stick-Stick-CSF	30.9	0.48	1.70				0.1	81.6	0.50	1.70	0.62	0.43		
H	Tensor-Stick-Stick-CSF-diff	13.5	0.40	0.98				0.1	82.5	0.58	2.69	0.51	0.33		
L	Ball-Stick-CSF-diff	69.2	0.60	2.12				0.1		0.39	0.83				
L	Ball-Stick-CSF	100.0	0.64	1.51				0.1		0.35	1.51				
M	Tensor-Stick-CSF	33.6	0.48	1.69				0.1		0.51	1.69	0.62	0.42		
H	Tensor-Stick-CSF-diff	16.9	0.40	0.99				0.1		0.59	2.63	0.51	0.33		
M	Ball-Cylinder-CSF-diff	23.0	0.80	1.68			4.1	0.1		0.20	1.43				
M	Ball-Cylinder-CSF	23.5	0.81	1.64			4.1	0.1		0.18	1.64				
M	MMWMD	23.0	0.63	1.70			3.5	0.1		0.36	1.70	0.61			
M	Tensor-Cylinder-CSF	13.0	0.59	1.67			6.1	0.1		0.41	1.67	0.57	0.30		
H	Tensor-Cylinder-CSF-diff	0.0	0.57	1.83			5.5	0.1		0.43	1.98	0.40	0.21		
M	Zeppelin-Zeppelin-CSF	35.3	0.50	1.68				0.0	89.9	0.49	1.68	0.37			
M	Zeppelin-Zeppelin-CSF-diff	20.8	0.52	2.33				0.1	89.0	0.47	1.45	0.13			
WM $T_2 = 51$ ms															

stable and repeatable parameter estimates that reflect the underlying tissue architecture. The stability requirement favours simple models with complexity similar to those in current microstructure imaging techniques, such as NODDI. Tailored experiment design (Alexander, 2008) and more reliable location of globally optimal parameter estimates (Daducci et al., 2015) may help stabilise parameter estimates from more complex models to some extent. However, it seems unlikely that PGSE acquisition protocols on current widely available scanners, such as the Achieva, will support much more complex models than those in currently popular techniques. The second broad conclusion is

that no single model clearly outperforms all the others. While most microstructure imaging techniques adopt a single model for all voxels, future techniques may benefit from retaining a range of models and making a selection at each voxel or by averaging predictions from multiple models in the spirit of Burnham and Anderson (2002) and explored to some extent by Stamm et al. (2014).

This study has a number of limitations and areas for further work. We study only a subset of the wide set of possible three compartment models, with the aim of investigating consistency at different gradient strengths rather than identifying the best from all possible models.

Table 3
The stability of parameters across 200 data sets (fifty bootstrap sets for each of the four ROI voxels). In addition to the mean estimate across all fits to the bootstrap data sets, we also report beside it, in blue superscript, the standard deviation as a percentage of the mean. More complex models generally have higher parameter variance.

		Cylinder/Bingham/Stick						Tensor/ Ball					
Models	SSE (0=best)	Vol.Fraction	Diffusivity ($\mu\text{m}^2/\text{s}$)	k_1	Radius (μm)	Vol.Fraction	Axial Diff. ($\mu\text{m}^2/\text{s}$)	Radial Diff. ($\mu\text{m}^2/\text{s}$)	Radial Diff. ($\mu\text{m}^2/\text{s}$)	Radial Diff. ($\mu\text{m}^2/\text{s}$)	Radial Diff. ($\mu\text{m}^2/\text{s}$)	Radial Diff. ($\mu\text{m}^2/\text{s}$)	Radial Diff. ($\mu\text{m}^2/\text{s}$)
H	Ball-Bingham-CSF-diff	1.5 ^{17.8}	0.74 ^{5.9}	2.62 ^{6.6}	15 ^{25.6}		0.23 ^{20.4}	0.70 ^{15.8}					
M	Ball-Bingham-CSF	35.2 ^{13.0}	0.85 ^{8.5}	2.22 ^{6.2}	10 ^{12.8}		0.12 ^{64.2}	2.22 ^{6.2}					
M	NODDI	14.6 ^{14.5}	0.64 ^{8.0}	1.90 ^{0.0}	11 ^{9.0}		0.34 ^{17.9}	1.90 ^{0.0}	0.65 ^{17.9}				
H	Tensor-Bingham-CSF	1.8 ^{17.3}	0.63 ^{7.3}	1.95 ^{4.2}	23 ^{27.2}		0.35 ^{13.1}	1.95 ^{4.2}	0.64 ^{25.2}	0.52 ^{29.4}			
H	Tensor-Bingham-CSF-diff	0.0 ^{17.5}	0.68 ^{9.0}	2.44 ^{17.8}	20 ^{58.6}		0.30 ^{18.9}	1.18 ^{45.5}	0.64 ^{26.0}	0.53 ^{29.7}			
L	Ball-Stick-Stick-CSF-diff	62.2 ^{6.7}	0.59 ^{6.8}	2.01 ^{3.9}			0.37 ^{15.0}	0.87 ^{16.0}					
L	Ball-Stick-Stick-CSF	92.5 ^{5.2}	0.63 ^{7.3}	1.53 ^{2.1}			0.32 ^{20.5}	1.53 ^{2.1}					
M	Tensor-Stick-Stick-CSF	20.9 ^{10.3}	0.45 ^{5.6}	1.67 ^{1.9}			0.51 ^{6.0}	1.67 ^{1.9}	0.61 ^{19.9}	0.43 ^{28.4}			
H	Tensor-Stick-Stick-CSF-diff	2.7 ^{16.0}	0.39 ^{6.0}	1.00 ^{3.6}			0.56 ^{5.6}	2.54 ^{5.7}	0.54 ^{19.2}	0.36 ^{27.8}			
L	Ball-Stick-CSF-diff	71.5 ^{6.0}	0.60 ^{6.8}	2.01 ^{3.4}			0.38 ^{12.8}	0.87 ^{15.5}					
L	Ball-Stick-CSF	100.0 ^{5.7}	0.64 ^{7.6}	1.50 ^{2.4}			0.33 ^{17.4}	1.50 ^{2.4}					
M	Tensor-Stick-CSF	24.0 ^{10.7}	0.45 ^{5.0}	1.65 ^{2.6}			0.52 ^{4.6}	1.65 ^{2.6}	0.62 ^{19.0}	0.42 ^{29.8}			
H	Tensor-Stick-CSF-diff	3.8 ^{16.3}	0.40 ^{5.3}	0.98 ^{3.3}			0.58 ^{4.3}	2.56 ^{5.2}	0.55 ^{17.6}	0.36 ^{29.5}			
M	Ball-Cylinder-CSF-diff	43.6 ^{9.5}	0.78 ^{11.5}	1.91 ^{7.0}	6.1 ^{9.6}		0.19 ^{50.3}	0.64 ^{80.1}					
M	Ball-Cylinder-CSF	49.3 ^{10.1}	0.85 ^{8.9}	1.61 ^{2.2}	6.1 ^{3.6}		0.13 ^{68.5}	1.61 ^{2.2}					
M	MMWMD	31.1 ^{7.2}	0.65 ^{14.7}	1.70 ^{0.0}	5.1 ^{13.7}		0.33 ^{32.3}	1.70 ^{0.0}	0.56 ^{32.3}				
M	Tensor-Cylinder-CSF	17.7 ^{9.0}	0.49 ^{16.2}	1.65 ^{2.7}	5.6 ^{48.1}		0.49 ^{15.9}	1.65 ^{2.7}	0.49 ^{57.0}	0.31 ^{72.8}			
H	Tensor-Cylinder-CSF-diff	0.4 ^{13.6}	0.47 ^{17.8}	1.10 ^{25.8}	4.0 ^{45.1}		0.50 ^{18.5}	2.68 ^{10.7}	0.52 ^{25.3}	0.31 ^{46.7}			
M	Zeppelin-Zeppelin-CSF	23.7 ^{7.9}	0.49 ^{36.6}	1.74 ^{6.0}			0.49 ^{35.9}	1.74 ^{6.0}	0.42 ^{127.2}				
M	Zeppelin-Zeppelin-CSF-diff	11.2 ^{13.3}	0.49 ^{14.6}	1.91 ^{41.4}			0.49 ^{15.2}	1.97 ^{40.6}	0.28 ^{105.1}				

H	Ball-Bingham-CSF-diff	10.7 ^{16.6}	0.75 ^{3.6}	2.57 ^{2.2}	16 ^{10.8}		0.24 ^{10.6}	0.81 ^{10.3}					
M	Ball-Bingham-CSF	39.6 ^{14.8}	0.83 ^{5.4}	2.14 ^{5.7}	13 ^{19.6}		0.16 ^{28.6}	2.14 ^{5.7}					
M	NODDI	18.5 ^{12.6}	0.63 ^{3.3}	1.90 ^{0.0}	13 ^{13.6}		0.37 ^{6.2}	1.90 ^{0.0}	0.69 ^{6.2}				
H	Tensor-Bingham-CSF	8.6 ^{16.5}	0.64 ^{5.6}	1.97 ^{2.9}	21 ^{12.8}		0.35 ^{9.2}	1.97 ^{2.9}	0.72 ^{15.8}	0.60 ^{16.8}			
H	Tensor-Bingham-CSF-diff	7.7 ^{16.4}	0.58 ^{9.0}	1.63 ^{10.7}	22 ^{17.9}		0.42 ^{11.6}	2.42 ^{9.7}	0.68 ^{16.3}	0.51 ^{19.3}			
L	Ball-Stick-Stick-CSF-diff	64.8 ^{6.2}	0.60 ^{3.8}	2.12 ^{5.6}			0.39 ^{7.4}	0.89 ^{10.3}					
L	Ball-Stick-Stick-CSF	93.9 ^{6.1}	0.63 ^{4.3}	1.55 ^{3.5}			0.35 ^{10.1}	1.55 ^{3.5}					
M	Tensor-Stick-Stick-CSF	28.3 ^{9.4}	0.48 ⁴	1.71 ^{3.0}			0.50 ^{3.8}	1.71 ^{3.0}	0.67 ^{15.2}	0.45 ^{13.8}			
H	Tensor-Stick-Stick-CSF-diff	9.4 ^{13.3}	0.40 ^{3.4}	1.00 ^{6.0}			0.58 ^{3.3}	2.71 ^{4.0}	0.54 ^{14.9}	0.35 ^{14.6}			
L	Ball-Stick-CSF-diff	68.5 ^{7.0}	0.60 ^{3.9}	2.12 ^{5.2}			0.39 ^{6.5}	0.88 ^{9.7}					
L	Ball-Stick-CSF	100.0 ^{7.5}	0.64 ^{4.4}	1.53 ^{3.0}			0.36 ^{8.8}	1.53 ^{3.0}					
M	Tensor-Stick-CSF	30.6 ^{10.4}	0.48 ^{3.5}	1.70 ^{2.6}			0.51 ^{3.2}	1.70 ^{2.6}	0.66 ^{15.0}	0.45 ^{14.0}			
H	Tensor-Stick-CSF-diff	12.6 ^{14.5}	0.41 ^{3.4}	1.01 ^{5.7}			0.59 ^{2.5}	2.64 ^{4.1}	0.55 ^{14.1}	0.36 ^{15.1}			
M	Ball-Cylinder-CSF-diff	22.8 ^{6.2}	0.77 ^{4.8}	1.70 ^{3.5}	3.9 ^{3.4}		0.22 ^{18.6}	1.42 ^{10.4}					
M	Ball-Cylinder-CSF	23.7 ^{6.2}	0.79 ^{4.9}	1.65 ^{3.2}	3.9 ^{3.6}		0.21 ^{21.0}	1.65 ^{3.2}					
M	MMWMD	21.5 ^{7.6}	0.61 ^{4.9}	1.70 ^{0.0}	3.3 ^{5.0}		0.38 ^{8.8}	1.70 ^{0.0}	0.65 ^{8.8}				
M	Tensor-Cylinder-CSF	12.1 ^{8.1}	0.60 ^{17.5}	1.69 ^{2.5}	5.3 ^{50.5}		0.39 ^{26.1}	1.69 ^{2.5}	0.79 ^{67.4}	0.46 ^{69.2}			
H	Tensor-Cylinder-CSF-diff	0.0 ^{11.8}	0.57 ^{9.9}	1.54 ^{37.4}	4.3 ^{50.8}		0.43 ^{13.7}	2.53 ^{29.0}	0.65 ^{53.4}	0.38 ^{57.2}			
M	Zeppelin-Zeppelin-CSF	32.3 ^{9.2}	0.50 ^{18.9}	1.70 ^{2.9}			0.49 ^{19.1}	1.70 ^{2.9}	0.39 ^{95.1}				
M	Zeppelin-Zeppelin-CSF-diff	16.9 ^{12.2}	0.49 ^{9.3}	1.85 ^{44.6}			0.50 ^{9.1}	1.96 ^{41.7}	0.28 ^{92.7}				

CSF Vol.Fract = 0.02

WM T₂ = 58.3

ACH-genu

CSF Vol.Fract = 0.01

WM T₂ = 51.8

CON-genu

Preliminary work suggests little benefit from obvious extensions of the set of models we present results from here, such as Watson or Bingham distributions of cylinders (Zhang et al., 2011b), GDR-cylinders (Assaf et al., 2008; Panagiotaki et al., 2012), or mixtures of Watson and Bingham distributions of sticks (Kaden et al., 2007; Sotiropoulos et al., 2012). While the results suggest that certain models provide less reliable estimates than others, we cannot conclude that the signal is not sensitive to the effects modelled by any compartment in particular. Because the data

we acquire are not optimised in the estimation of any specific parameter we can only draw general conclusions about model compartments.

Though we here report *AIC*, as suggested in Burnham and Anderson (2002), we also checked the *BIC* results for comparison, which agree strongly with the *AIC*. But we expect that *AIC* and *BIC* differences would become significant for smaller and sparser data sets. It could well be that none of these models are self-sufficient enough in describing the data – to quote the statistician G.E.P. Box, “all models are wrong, but

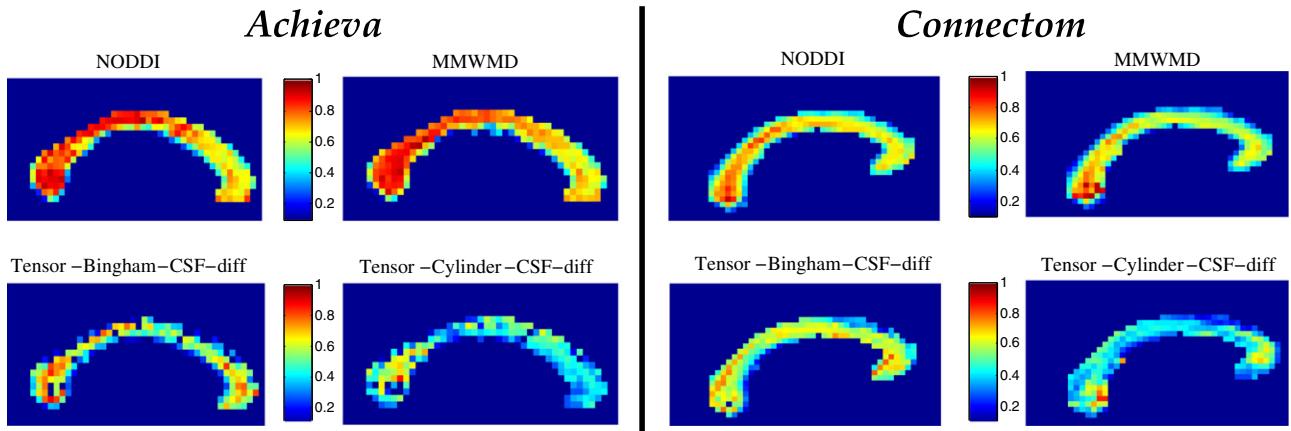


Fig. 9. Intracellular volume fraction maps of four representative models fitted to the whole corpus callosum *Achieva* data set (left) and *Connectom* data set (right). NODDI and MMWMD are simplified constructions of Tensor-Bingham-CSF-diff and Tensor-Cylinder-CSF-diff, respectively.

some are useful” – but they nonetheless provide us with useful and interesting statistics, which no doubt will inform even better models in the future. In this study we start with a simple and established family. However, other better models may exist, and our data will be available online for others to compare other variants (for more information go to the ‘Tutorials/Compartment Models’ section on the Camino website <http://cmic.cs.ucl.ac.uk/camino/>). One possibility is to enrich each model with effects such as compartmental exchange/permeability or time-dependent diffusivity (Novikov et al., 2014). Moreover, such models may well provide benefits in other regions of the brain with more complex fibre architecture; possibly by also exploiting measurements from other pulse sequences, such as Stimulated-Echo Acquisition Mode (STEAM) (Merboldt et al., 1969; Tanner, 1970), Oscillating-Gradient Spin-Echo (OGSE) (Callaghan and Stepišnik, 1995), or Double Pulsed-Field Gradient Sequence (dPFG) sequences (Mittra, 1995; Cheng and Cory, 1999; Komlosh et al., 2007).

Artefacts in the experimental acquisition protocol can potentially affect the results, so we focus only on the mid-sagittal corpus callosum and, furthermore, mostly on just the genu, since our highly experimental acquisition produced the most usable data in that region. We cannot rule out that hidden artefacts affect the results in this area too, although we believe their influence to be minor. Future work with optimised pulse sequences will help eliminate such affects, and also enable a comparison of results to determine consistency across multiple subjects. Other minor effects may also influence results. For example, we do not correct for alterations to the B-matrix (Leemans and Jones, 2009) from movement, although those corrections are minor and are unlikely to have a major effect. Similarly, we account for non-central chi noise bias in only a simple way via the offset-Gaussian fitting objective function and more sophisticated techniques, e.g. from Veraart et al. (2013), might be beneficial when extrapolating from the results in this paper to more complete microstructure imaging techniques. Although the other areas of the corpus callosum appear to give similar results, more peripheral WM with greater orientation dispersion and prevalent crossings are likely to yield different conclusions. For example, Table 2 shows that in fact the Bingham distributions from the genu are quite isotropic so a Watson distribution would probably be sufficient there; however, in regions of fanning fibres the anisotropy of the Bingham distribution is likely to be more important (Tariq et al., 2014). Other pulse sequences may also provide access to other parameters that the models we use do not consider. For example, STEAM and Pulsed-Gradient Stimulated-Echo (PGSTE) may inform on WM exchange rates (as in Nedjati-Gilani et al. (2014) and Nilsson et al. (2009), respectively) and dPFG may inform on fibre curvature (Jespersen and Buhl, 2011).

One final point to highlight is that we include T_2 parameters in models for the diffusion MRI signal for the first time in an extensive model selection experiment. The T_2 parameters are essential to explain

the signal well and careful estimation of their values is important, as even quite small errors can have a significant influence on the estimation of the other parameters. Additional experiments in Ferizi (2014) using the same data sets we report here study these effects in more detail.

Acknowledgments

EPSRC support this work through grants EP/G007748, EP/L022680/1, and EP/I027084/01.

The NMR unit is supported by the MS Society of Great Britain (grant 892/08), the UK’s National Institute of Health Research (Capital Project number R&D03/10/RAG0449) and UCLH.

We thank Dr José G. Raya, of the Department of Radiology, New York University School of Medicine, for the support he has given to the first author (U.F.) during the review stage of this article.

Appendix A

The normalised signal, as a result of diffusion along all directions \mathbf{n} , for an intracellular or extracellular compartment follows (Zhang and Alexander, 2010)

$$A = \int_{S^2} f(\mathbf{n}) A_{comp} d\mathbf{n} \quad (4)$$

where the integrand’s models for f and A_{comp} vary across the compartments, each of which we will explain separately below.

Distribution f

In the case of Bingham-distributed sticks, $f = \exp[\kappa_1(\boldsymbol{\mu}_1 \cdot \mathbf{n}) + \kappa_2(\boldsymbol{\mu}_2 \cdot \mathbf{n})]/F(\kappa_1, \kappa_2)$, where F is a confluent hypergeometric function, κ_1 and κ_2 are the eigenvalues (concentration parameters) with respect to each eigenvector $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ of the Bingham distribution (Sotiropoulos et al., 2012; Tariq et al., 2014). For all other compartments, f is a Dirac delta function.

Model for A_{comp}

The Tensor model has $A_{comp} = \exp(-b\mathbf{q}^T \mathbf{D} \mathbf{q})$, where $\mathbf{D} = d_{\parallel} \mathbf{nn}^T + d_{\perp 1} \mathbf{n}_{\perp 1} \mathbf{n}_{\perp 1}^T + d_{\perp 2} \mathbf{n}_{\perp 2} \mathbf{n}_{\perp 2}^T$ is the diagonalised Tensor matrix with eigenvalues $\{d_{\parallel}, d_{\perp 1}, d_{\perp 2}\}$ and their respective eigenvectors $\{\mathbf{n}, \mathbf{n}_{\perp 1}, \mathbf{n}_{\perp 2}\}$ and \mathbf{q} as the normalised wave vector. As special cases of the above Tensor, the isotropic Ball has $d_{\parallel} = d_{\perp 1} = d_{\perp 2}$, whereas the anisotropic (one-

dimensional) Stick, which is used to model the single Stick, two Sticks or Bingham-distributed sticks, has $d_{\perp 1} = 0 = d_{\perp 2}$.

The Cylinder model uses the Gaussian phase distribution approximation (Murday and Cotts, 1968; Stepišnik, 1993; Vangelder et al., 1994; Alexander, 2008). This restricted signal is a product of the parallel A_{\parallel} and radial A_{\perp} , both being functions of applied diffusion gradient of magnitude G , with radial component G_{\perp} and axial G_{\parallel} relative to the cylinder axis. The axial signal is:

$$A_{\parallel}(G_{\parallel}, \Delta, \delta) = e^{-bd_1} = e^{-(\Delta-\delta/3)\gamma^2\delta^2 G_{\parallel}^2 d_{\parallel}} \quad (5)$$

whereas the radial signal component is:

$$A_{\perp}(G_{\perp}, \Delta, \delta) = \exp\left(-2\gamma^2 G_{\perp}^2 \sum_{m=1}^{\infty} \frac{2d_{\parallel}\alpha_m^2\delta - 2 + 2e^{-d_{\parallel}\alpha_m^2\delta} + 2e^{-d_{\parallel}\alpha_m^2\Delta} - e^{-d_{\parallel}\alpha_m^2(\Delta-\delta)} - e^{-d_{\parallel}\alpha_m^2(\Delta+\delta)}}{d_{\parallel}^2\alpha_m^6(R^2\alpha_m^2 - 1)}\right) \quad (6)$$

where R is the cylinder radius, d_{\parallel} and d_{\perp} are the apparent axial and radial diffusion coefficients, and $\alpha_m R$ is the m -th zero of the derivative of the Bessel function of the first kind, order one.

References

- Aboitiz, F., Scheibel, A.B., Fisher, R.S., Zaidel, E., 1992. Fiber composition of the human corpus callosum. *Brain Res.* 598, 143–153.
- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Control* 19, 716–723.
- Alexander, D.C., 2008. A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. *Magn. Reson. Med.* 60, 439–448.
- Alexander, D.C., Hubbard, P.L., Hall, M.G., Moore, E.A., Ptito, M., Parker, G.J., Dyrby, T.B., 2010. Orientationally invariant indices of axon diameter and density from diffusion MRI. *NeuroImage* 52, 1374–1389.
- Assaf, Y., Basser, P.J., 2005. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. *NeuroImage* 27, 48–58.
- Assaf, Y., Freidlin, R.Z., Rohde, G.K., Basser, P.J., 2004. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magn. Reson. Med.* 52, 965–978.
- Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y., Basser, P.J., 2008. Axcaliber: a method for measuring axon diameter distribution from diffusion MRI. *Magn. Reson. Med.* 59, 1347–1354.
- Behrens, T., Woolrich, M., Jenkinson, M., Johansen-Berg, H., Nunes, R., Clare, S., Matthews, P., Brady, J., Smith, S., 2003. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn. Reson. Med.* 50, 1077–1088.
- Behrens, T., Berg, H.J., Jbabdi, S., Rushworth, M., Woolrich, M., 2007. Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? *NeuroImage* 34, 144–155.
- Burnham, K.P., Anderson, D.R., 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer.
- Callaghan, P.T., Stepišnik, J., 1995. Frequency-domain analysis of spin motion using modulated-gradient NMR. *J. Magn. Reson. Ser. A* 117, 118–122.
- Cheng, Y., Cory, D.G., 1999. Multiple scattering by NMR. *J. Am. Chem. Soc.* 121, 7935–7936.
- Daducci, A., Canales-Rodriguez, E.J., Zhang, H., Dyrby, T.B., Alexander, D.C., Thiran, J.P., 2015. Accelerated microstructure imaging via convex optimization (AMICO) from diffusion MRI data. *NeuroImage* 105, 32–44.
- Duval, T., McNab, J., Setsompop, K., Witzel, T., Schneider, T., Susie Yi, H., Keil, B., Klawitter, E., Wald, L., Cohen-Adad, J., 2014. In vivo estimation of axon diameter in the human spinal cord using 300 mT/m gradients. *Joint Annual Meeting ISMRM-ESMRMB 2014*, Milan, Italy.
- Dyrby, T.B., Hall, M.G., Ptito, M., Alexander, D., et al., 2013. Contrast and stability of the axon diameter index from microstructure imaging with diffusion MRI. *Magn. Reson. Med.* 70, 711–721.
- Efron, B., 1979. Bootstrap methods: another look at the jackknife. *Ann. Stat.* 1–26.
- Ferizi, U., 2014. *Compartment Models and Model Selection for In-Vivo Diffusion-MRI of Human Brain White Matter* Ph.D. thesis. UCL (University College London).
- Ferizi, U., Schneider, T., Tariq, M., Wheeler-Kingshott, C., Zhang, H., Alexander, D., 2013. The importance of being dispersed: a ranking of diffusion MRI models for fibre dispersion using in vivo human brain data. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013 Lecture Notes in Computer Science*. Springer, Berlin Heidelberg, pp. 74–81.
- Ferizi, U., Schneider, T., Panagiotaki, E., Nedjati-Gilani, G., Zhang, H., Wheeler-Kingshott, C.A., Alexander, D.C., 2014. A ranking of diffusion MRI compartment models with in vivo human brain data. *Magn. Reson. Med.* 72, 1785–1792.
- Gudbjartsson, H., Patz, S., 1995. The Rician distribution of noisy MRI data. *Magn. Reson. Med.* 34, 910–914.
- Highley, J.R., Esiri, M.M., McDonald, B., Cortina-Borja, M., Herron, B.M., Crow, T.J., 1999. The size and fibre composition of the corpus callosum with respect to gender and schizophrenia: a post-mortem study. *Brain* 122, 99–110.
- Hosey, T., Williams, G., Ansoorge, R., 2005. Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magn. Reson. Med.* 54, 1480–1489.
- Huang, S.Y., Nimmennmaa, A., Witzel, T., Duval, T., Cohen-Adad, J., Wald, L.L., McNab, J.A., 2015. The impact of gradient strength on in vivo diffusion MRI estimates of axon diameter. *NeuroImage* 106, 464–472.
- Jbabdi, S., Sotiropoulos, S., Savio, A., Graa, M., Behrens, T., 2012. Model-based analysis of multishell diffusion MR data for tractography: how to get over fitting problems. *Magn. Reson. Med.* 68, 1846–1855.
- Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* 17, 825–841.
- Jespersen, S.N., Buhl, N., 2011. The displacement correlation tensor: microstructure, ensemble anisotropy and curving fibers. *J. Magn. Reson.* 208, 34–43.
- Jeurissen, B., Leemans, A., Tournier, J.D., Jones, D.K., Sijbers, J., 2013. Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Hum. Brain Mapp.* 34, 2747–2766.
- Kaden, E., Knösche, T.R., Anwander, A., 2007. Parametric spherical deconvolution: inferring anatomical connectivity using diffusion MR imaging. *NeuroImage* 37, 474–488.
- Komlosch, M., Horkay, F., Freidlin, R., Nevo, U., Assaf, Y., Basser, P., 2007. Detection of microscopic anisotropy in gray matter and in a novel tissue phantom using double pulsed gradient spin echo MR. *J. Magn. Reson.* 189, 38–45.
- Kunz, N., Zhang, H., Vasung, L., O'Brien, K.R., Assaf, Y., Lazeyras, F., Alexander, D.C., Hüppi, P.S., 2014. Assessing white matter microstructure of the newborn with multi-shell diffusion MRI and biophysical compartment models. *NeuroImage* 96, 288–299.
- Leemans, A., Jones, D.K., 2009. The b-matrix must be rotated when correcting for subject motion in DTI data. *Magn. Reson. Med.* 61, 1336–1349.
- MacKay, A., Whittall, K., Adler, J., Li, D., Paty, D., Graeb, D., 1994. In vivo visualization of myelin water in brain by magnetic resonance. *Magn. Reson. Med.* 31, 673–677. <http://dx.doi.org/10.1002/mrm.1910310614>.
- McNab, J.A., Edlow, B.L., Witzel, T., Huang, S.Y., Bhat, H., Heberlein, K., Feiweier, T., Liu, K., Keil, B., Cohen-Adad, J., et al., 2013. The human connectome project and beyond: initial applications of 300 mT/m gradients. *NeuroImage* 80, 234–245.
- Merboldt, K.D., Hanicke, W., Frahm, J., 1969. Self-diffusion NMR imaging using stimulated echoes. *J. Magn. Reson.* 64, 479–486.
- Mitra, P.P., 1995. Multiple wave-vector extensions of the NMR pulsed-field-gradient spin-echo diffusion measurement. *Phys. Rev. B* 51, 15074.
- Murday, J., Cotts, R.M., 1968. Self-diffusion coefficient of liquid lithium. *J. Chem. Phys.* 48, 4938–4945.
- Nedjati-Gilani, G.L., Schneider, T., Hall, M.G., Wheeler-Kingshott, C.A., Alexander, D.C., 2014. Machine learning based compartment models with permeability for white matter microstructure imaging. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2014*. Springer, pp. 257–264.
- Nilsson, M., Lätt, J., Nordh, E., Wirestam, R., Ståhlberg, F., Brockstedt, S., 2009. On the effects of a varied diffusion time in vivo: is the diffusion in white matter restricted? *Magn. Reson. Imaging* 27, 176–187.
- Novikov, D.S., Jensen, J.H., Helpert, J.A., Fieremans, E., 2014. Revealing mesoscopic structural universality with diffusion. *Proc. Natl. Acad. Sci.* 111, 5088–5093.
- Panagiotaki, E., Schneider, T., Siow, B., Hall, M.G., Lythgoe, M.F., Alexander, D.C., 2012. Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison. *NeuroImage* 59, 2241–2254.
- Parker, G.J., Alexander, D.C., 2003. Probabilistic Monte Carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. *Information Processing in Medical Imaging*. Springer, pp. 684–695.
- Richardson, S., Siow, B., Panagiotaki, E., Schneider, T., Lythgoe, M.F., Alexander, D.C., 2013. Viable and fixed white matter: diffusion magnetic resonance comparisons and contrasts at physiological temperature. *Magn. Reson. Med.* 72, 1151–1161.
- Scherrer, B., Schwartzman, A., Taquet, M., Prabhu, S.P., Sahin, M., Akhondi-Asl, A., Warfield, S.K., 2013. Characterizing the distribution of anisotropic micro-structural environments with diffusion-weighted imaging (DIAMOND). *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*. Springer, pp. 518–526.
- Setsompop, K., Kimmlingen, R., Eberlein, E., Witzel, T., Cohen-Adad, J., McNab, J.A., Keil, B., Tisdall, M.D., Hoect, P., Dietz, P., et al., 2013. Pushing the limits of in vivo diffusion MRI for the human connectome project. *NeuroImage* 80, 220–233.
- Seunarine, K.K., Alexander, D.C., 2009. Multiple fibers: beyond the diffusion tensor. In: Johansen-Berg, H., Behrens, T.E. (Eds.), *Diffusion MRI: From Quantitative Measurement to In Vivo Neuroanatomy* (Chapter 4). Academic Press.
- Sotiropoulos, S.N., Behrens, T.E., Jbabdi, S., 2012. Ball and rackets: inferring fiber fanning from diffusion-weighted MRI. *NeuroImage* 60, 1412–1425.
- Stamm, A., Scherrer, B., Commowick, O., Barillot, C., Warfield, S.K., et al., 2014. Fast and robust detection of the optimal number of fascicles in diffusion images using model averaging theory. *ISMRM Proceedings*.
- Stanisz, G.J., Wright, G.A., Henkelman, R.M., Zafer, A., 1997. An analytical model of restricted diffusion in bovine optic nerve. *Magn. Reson. Med.* 37, 103–111.
- Stepišnik, J., 1993. Time-dependent self-diffusion by NMR spin-echo. *Phys. B Condens. Matter* 183, 343–350.
- Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc. Ser. B Methodol.* 111–147.
- Tanner, J., 1970. Use of the stimulated echo in NMR diffusion studies. *J. Chem. Phys.* 52, 2523–2526.
- Tanner, J., Stejskal, E.O., 1968. Restricted self-diffusion of protons in colloidal systems by the pulsed-gradient, spin-echo method. *J. Chem. Phys.* 49, 1768–1777.
- Tariq, M., Schneider, T., Alexander, D.C., Wheeler-Kingshott, C.A., Zhang, H., 2014. In vivo estimation of dispersion anisotropy of neurites using diffusion MRI. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2014*. Springer, pp. 241–248.

- Tournier, J.D., Calamante, F., Connelly, A., 2007. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *NeuroImage* 35, 1459–1472.
- Tournier, J.D., Mori, S., Leemans, A., 2011. Diffusion tensor imaging and beyond. *Magn. Reson. Med.* 65, 1532–1556.
- Tuch, D.S., Reese, T.G., Wiegell, M.R., Makris, N., Belliveau, J.W., Wedeen, V.J., 2002. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn. Reson. Med.* 48, 577–582.
- Vangelder, P., DesPres, D., Vanzijl, P., Moonen, C., 1994. Evaluation of restricted diffusion in cylinders. Phosphocreatine in rabbit leg muscle. *J. Magn. Reson. B* 103, 255–260.
- Veraart, J., Rajan, J., Peeters, R.R., Leemans, A., Sunaert, S., Sijbers, J., 2013. Comprehensive framework for accurate diffusion MRI parameter estimation. *Magn. Reson. Med.* 70, 972–984.
- Whittall, K.P., Mackay, A.L., Graeb, D.A., Nugent, R.A., Li, D.K., Paty, D.W., 1997. In vivo measurement of T2 distributions and water contents in normal human brain. *Magn. Reson. Med.* 37, 34–43.
- Winston, G.P., Micallef, C., Symms, M.R., Alexander, D.C., Duncan, J.S., Zhang, H., 2014. Advanced diffusion imaging sequences could aid assessing patients with focal cortical dysplasia and epilepsy. *Epilepsy Res.* 108, 336–339.
- Zhang, H., Alexander, D., 2010. Axon diameter mapping in the presence of orientation dispersion with diffusion MRI. In: Jiang, T., Navab, N., Pluim, J.P., Viergever, M.A. (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010 Lecture Notes in Computer Science* volume 6361. Springer, Berlin Heidelberg, pp. 640–647. http://dx.doi.org/10.1007/978-3-642-15705-9_78.
- Zhang, H., Dyrby, T.B., Alexander, D.C., 2011a. Axon diameter mapping in crossing fibers with diffusion MRI. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2011*. Springer, pp. 82–89.
- Zhang, H., Hubbard, P.L., Parker, G.J., Alexander, D.C., 2011b. Axon diameter mapping in the presence of orientation dispersion with diffusion MRI. *NeuroImage* 56, 1301–1315.
- Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C., 2012. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage* 61, 1000–1016.